2019

Application of mass spectrometric technologies to decipher high spatial resolution insights into plant metabolism

Liza Esther Alexander
Iowa State University

Application of mass spectrometric technologies to decipher high spatial resolution insights into plant metabolism

by

Liza Esther Alexander

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Molecular, Cellular and Developmental Biology

Program of Study Committee:
Basil J. Nikolau, Major Professor
Eve S. Wurtele
Harry T. Horner
Philip W. Becraft
Zengyi Shao

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2019

Dedicated to my love, Dinu, and my parents, Alexander Jacob and Elizabeth Alexander for their unconditional love and encouragement.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Basil J. Nikolau, for his support and guidance throughout the course of this research. I am grateful for the opportunity to work on exciting and challenging projects. I gained valuable exposure while at his lab through undergraduate mentoring programs, industrial internships, faculty training programs, and collaborative projects within Iowa State University (ISU) and internationally with RIKEN, Japan; these experiences have helped me grow tremendously as a researcher and as a person.

Next, I would like to thank my committee members, Dr. Eve S. Wurtele for the PMR database and Meta-Venn tool, Dr. Harry T. Horner for guidance on microscopy, Dr. Philip W. Becraft for his help and guidance in growing transgenic maize, and Dr. Zengyi Shao for guidance on faculty life as a "preparing future faculty" (PFF) mentor. I would also like to thank Dr. Marna D. Yandeau-Nelson, for the guidance on maize genetics, the opportunity to be part of her writing group, and for being an amazing PFF mentor, I will always cherish our conversations. I am grateful to my collaborators at ISU, Drs. Young-Jin Lee, R.S Houk, Adam D. Feenstra, Maria E. Duenas, Katherine-Jo Galayda, and Patrick McVey for the opportunity to learn mass spectral imaging. The second, third and fourth chapters of this dissertation would not have been possible without their collaboration. I would also like to thank my collaborators at RIKEN, Japan, Dr. Kazuki Saito, and Dr. Yozo Okazaki, for the cordial relationship maintained through the NSF-EAPSI program and for teaching me the lipidomics platform.

I am grateful to all my colleagues (past and present) in the Nikolau lab for their friendship, guidance, support, and encouragement during my graduate journey. I am thankful to Naazneen Sofeo, Kiran-Kumar Shivaiah, Xinyu Fu, Kenna Stenback, Elizabeth Chatt, Sara Hazinia, Troy Bunch, Geng Ding, Bri Vidrine, Keting Chen, Jennifer Chmielowski, Kayla

Flyckt, Lisa Vaknin, and Kathy Wiederin. I would like to particularly thank Dr. Libuse Brachova, our lab manager, for all the help, encouragement and advice. I specially thank, Dr. Ludmila Rizhsky for training me on techniques required for the epicuticular lipid project, and Dr. Bo Xie for the help on mass spectral imaging of Arabidopsis flowers. I would not have been able to handle all these projects without the help of my amazing undergraduate assistants, for which I am ever grateful to Katelyn Campbell, Jena Gilbertson, Michael Schelling, and Aeriel Davis.

I would like to acknowledge Drs. Ann Perera, Lucas Showman and Kirthi Narayanaswamy at the W.M Keck Metabolomics Research Laboratory for their assistance in operating GC instruments, Dr. Zhihong Song for training me on AMDIS data analysis, Tracey Stewart and Randell Den Adel of the Roy J. Carver High Resolution Microscopy Facility for their assistance in microscopy, and Dr. Charles Stewart from the Macromolecular X-ray Crystallography Facility for guidance on thermal shift assays.

Last, but certainly not least, I am forever grateful to my husband, Dinu Mathai, for his love, never ending support, and constant encouragement; I couldn't have done it without him! I am equally indebted to my loving family, my parents, Alexander Jacob and Elizabeth Alexander, my amazing brothers Sam J. Alexander, Steve M. Alexander, and John Alexander, Raisa Alexander, and the Mathai family, for their unconditional love, positive words, and for always being there by my side. Above all I thank God for being my pillar and strength throughout this journey.

ABSTRACT

Our understanding of how metabolic networks are structured and regulated is limited by the multicellular nature of plants. This complexity can be partly addressed through transcriptomics, proteomics, and metabolomics. However, these –omics platforms are limited by the dynamic nature of the metabolic networks, and by the techniques administered which utilize homogenized cells/tissues thereby losing spatial information. This limitation can be partially overcome by the use of more sophisticated analytical technologies, such as mass spectrometry imaging (MSI). This dissertation applies high-resolution MSI, using matrix assisted laser desorption ionization (MALDI) and electrospray laser desorption ionization (ELDI), as tools to decipher metabolite trafficking in plant cells inorder to elucidate complex metabolic processes that are distributed among different cell types and different cellular compartments.

MSI coupled with other analytical mass spectrometry technologies; gas- and liquid-chromatography (GC-/ LC-), is used to address four biological applications pertaining to maize germination, C4 photosynthesis, anthocyanin distribution in response to external stimuli, and epicuticular lipid accumulation in developing Arabidopsis flowers. Furthermore, this dissertation uses transgenic approaches (in Arabidopsis and yeast) together with mass spectrometry technologies to functionally characterize the role of two maize genes; *Glossy2* and a novel *Glossy2-like*, in epicuticular lipid biosynthesis.

CHAPTER 1: GENERAL INTRODUCTION

Background and Significance

The multicellular nature of plants has provided challenges to deciphering the biochemical mechanisms that regulate metabolic processes essential for plant survival. These metabolic networks are usually inferred by the asymmetric distribution of enzymes or mRNAs corresponding to individual components of the network. Furthermore, metabolites which are the fingerprints of enzyme catalyzed anabolic and catabolic cellular processes, can be distributed at sites different from their production. For example, the asymmetric distribution of C4 photosynthetic photoautotrophy is regulated by cellular compartmentalization of carbon assimilation in the mesophyll (M) cells, and carbon reduction in the bundle sheath (BS) (Edwards et al., 2001; Hatch, 1987; Nelson and Langdale, 1989; Sage and Monson, 1999). Additionally, studies have demonstrated that in maize leaf the distribution of benzoxazinoids, 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside (DIMBOA-Glc) and 2-hydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside (HMBOA-Glc), which serve as defense metabolites against herbivores and insects (Frey et al., 1997; Glauser et al., 2011; Meihls et al., 2013), are not only localized in M cells but are more specifically located in the M cells between each pair of vascular bundles (Korte et al., 2015). Another example, is the lesser-characterized metabolic trafficking and assembly of lipid molecules (Moore, 1993; Xu et al., 2008). These metabolites are produced by enzymatic reactions divided among different compartments at both the cellular and subcellular levels (Benning, 2008; Post-Beittenmiller, 1996).

These technical barriers can be partially overcome through advances in the field of genomics, transcriptomics, proteomics, and metabolomics. Forward genetic approaches have been

employed to identify and characterize genes involved in metabolic networks, however, a single genetic alteration may not be limited to one biochemical pathway (Koornneef et al., 1989; Schnable et al., 1994; Vogg et al., 2004). Additionally, the techniques applied in proteomics and metabolomics are generally performed using ground tissue extracts thereby losing *in situ* localization information. Technologies however do exist that provide spatial and temporal resolution of genes encoding mRNA and subsequent protein, these utilize target and reporter interactions (eg:- nucleic acid hybridization, antibodies) (Griffin et al., 1998; Küpper et al., 2007; McFadden, 1995), or gene and reporter molecule interactions (eg:- promoter-reporter or -protein fusions to GFP and luciferase) (Chalfie et al., 1994; Gallagher, 1992; Koo et al., 2007; Ow et al., 1986). However, the products measured by these -omics platforms are not necessarily dynamic. This gap in the biochemical characterization pertaining to the nature and regulation of metabolism as they move through metabolic networks in and among cells can be overcome by the use of more assiduous analytical technologies, namely mass spectrometry imaging (MSI) (McDonnell and Heeren, 2007).

Mass Spectrometry Imaging

MSI is a rapidly growing technology known for its ability to provide high spatial resolution, detect extremely low amounts of metabolites (sensitivity), and its chemical versatility. These properties have made MSI a desirable tool for analyzing spatial distributions of a variety of compounds such as lipids, proteins, and small molecules directly on or within plant and animal tissues (Amstalden van Hove et al., 2011; Angel and Caprioli, 2013; Jungmann and Heeren, 2012; Lee et al., 2012; Svatoš, 2010).

In a typical MS imaging experiment, the tissue specimen is mounted on plates, ablated by either a laser (UV or IR), ion beam (ESI capillary), or other ablation sources, and the resulting ions

collected at each point of ablation are then extracted into the mass spectrometer. A mass spectrometer consists of an ion source, detector, and mass analyzer. The chemical constituents detected by mass spectrometer are sorted based on the mass-to-charge ratio. Following data acquisition, the spatial distribution of ions of interest can be visualized based on the x-y coordinates of the collected spectra (Figure 1).

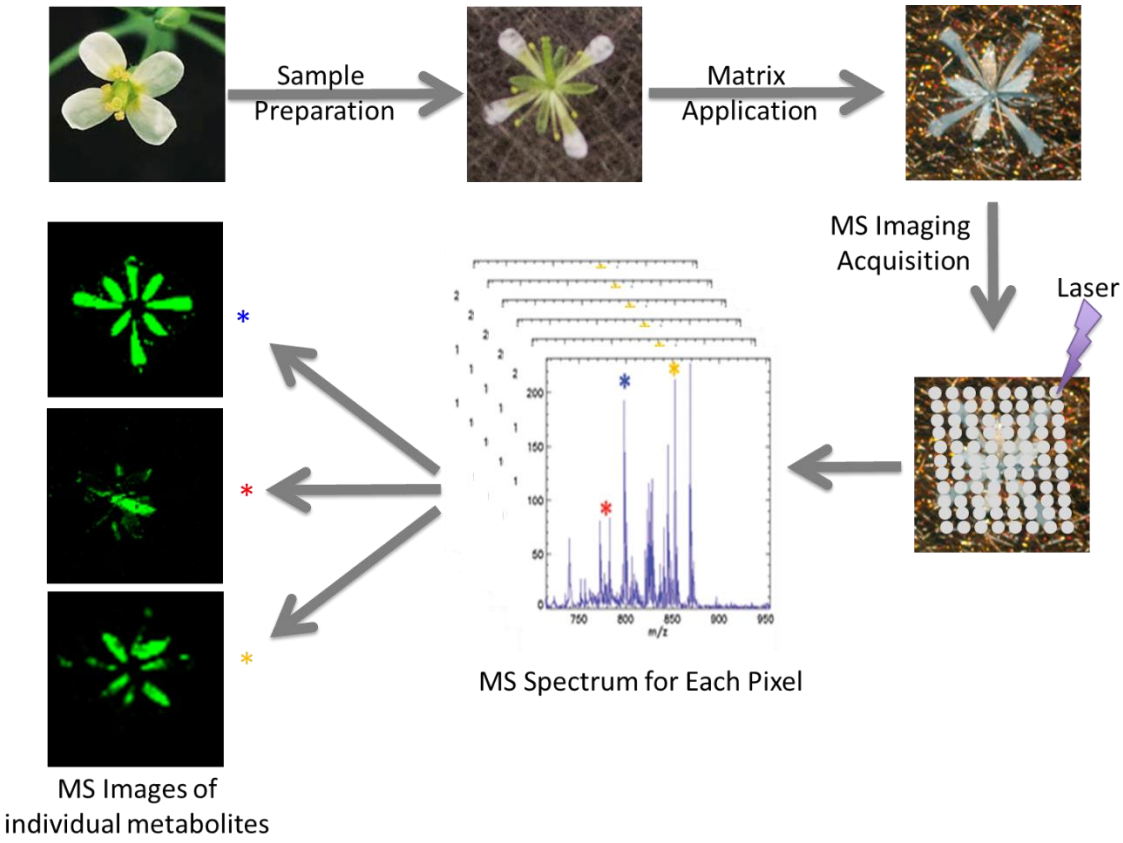


Figure 1. Schematic representation of mass spectral imaging. Samples are mounted on a plate and an appropriate ablation source is rastered across the sample. Each raster point denoted by the white dots generates a corresponding mass spectrum. The individual spectra are combined and mass spectral images are generated for each ion (as shown in green). For example, images in green are representative of the distribution of epicuticular lipids on the surface of Arabidopsis flowers.

The differences between MS imaging techniques can be classified based on how ions are generated, this process can either be conducted in ambient air (at atmospheric pressure) or by introduction into a vacuum (reduced pressure) chamber of the mass spectrometer. While there are several MS imaging techniques (for review, see Boughton et al., 2016; Lee et al., 2012) this

dissertation focuses specifically on the biological applications using matrix assisted laser desorption ionization (MALDI) and electrospray laser desorption ionization ELDI (Figure 2).

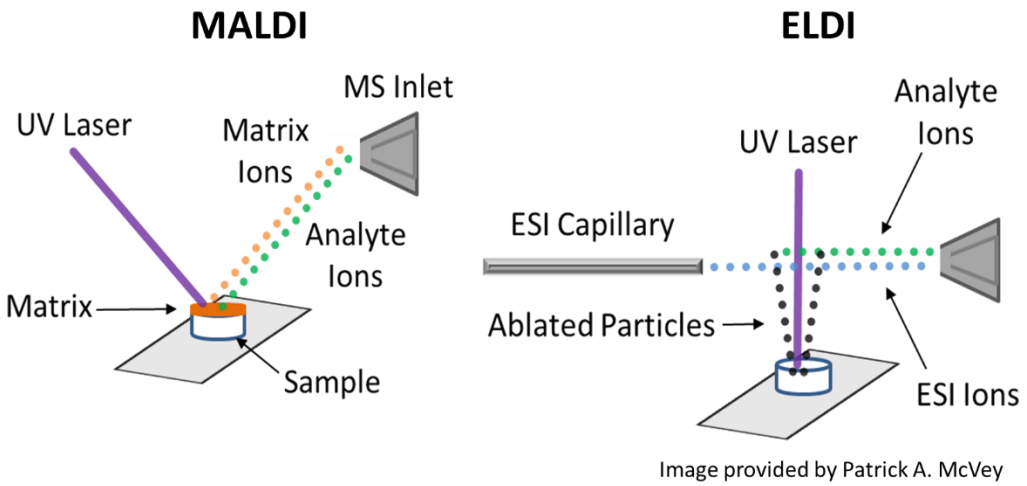


Figure 2. Schematic representation of MSI technologies used in the dissertation. In MALDI, samples are coated by UV-absorbing matrix. Both analyte and matrix ions enter the MS inlet to be detected by the mass spectrometer. In ELDI, sample is ablated using an UV-laser. The ablated particles are coupled with ions from the electrospray ionization (ESI) capillary. Both analyte ions and ESI ions are detected by the mass spectrometer.

Matrix assisted laser desorption ionization (MALDI) is the most widely used MSI technique in biology and can obtain spatial resolution as high as ~1 µm (Kompauer et al., 2017; Zavalin et al., 2015). MALDI is performed under reduced pressure (vacuum), and involves coupling the sample with different applied matrices which allows for the detection of a variety of metabolites (Klein et al., 2015; Shroff et al., 2015). Electrospray laser desorption ionization (ELDI) is conducted in ambient air, and combines laser ablation with electrospray ionization (ESI) (Jentaie et al., 2005). This technique is analogous to laser absorption electrospray ionization (LAESI) which uses an infrared laser instead of UV (Nemes and Vertes, 2007). This difference impacts the analytes extraction process which utilizes the pseudo-matrix properties of UV-absorbing compounds in the sample, rather than water as in IR-based methods. Additionally, in comparison to LAESI, the shorter inherent wavelength of UV lasers increases the ability of ELDI to achieve lower spot sizes and increases the spatial resolution. The biological application of ELDI

although not extended to plant applications has however been reported in two dry fungal species (i.e., *Ganoderma lucidum* and *Antrodia camphorate*) (Huang et al., 2012).

In this dissertation, MALDI and ELDI are used to demonstrate the capabilities of this technology in the biological context of 1) metabolic processes pertaining to maize seed germination, 2) exploring the genetic phenomenon of hybrid vigor in the context of membrane lipid distribution in photosynthetic cells of maize leaf, 3) the visual sectoring of light dependency on anthocyanin rich coleus leaves, and 4) epicuticular lipid accumulation in developing Arabidopsis flowers that express maize transgenes *Glossy2* and *Glossy2-like.* MSI coupled with LC and GC-MS analysis is used to provide insights to global metabolomics profiles. This integrated approach provides unprecedented insights into the spatially resolved "sharing" of metabolic processes that are sequestered among different cells, and tissues types. In addition, this dissertation explores the role of maize *Glossy2* gene family in epicuticular lipid biogenesis.

Epicuticular lipid biogenesis

Aerial organs of all terrestrial plants are covered with a thin protective hydrophobic lipid-based structure called the cuticle. The extracellular cuticular lipids are uniquely comprised of a mixture of very-long chain fatty acids (VLCFAs; up to 32-carbons in chain length) and their derivatives which include hydrocarbons, free fatty acids, ketones, alcohols, aldehydes and wax esters, and other terpene-type specialized metabolites (Kolattukudy, 1976; Martin and Juniper, 1970; Tulloch, 1976). These lipids are biosynthesized by a single layer of cells (the epidermal cells) and are unidirectionally secreted to the surface of the plant (Jellings and Leech, 1982; Jetter et al., 2000; Kolattukudy, 1965). Therefore, epidermal cells have unique attributes to balance the assembly of intracellular lipid components of cellular membranes, and achieve the biosynthesis and secretion of these extracellular lipids. Physiologically this process is genetically programmed,

but can change in response to environmental pressures (e.g., drought, temperature, pathogens) (Jenks et al., 1994; Kerstiens, 1996; Kolattukudy, 1985; Long, 2003; Riederer, 2006), making them important in agricultural crop development and crop yields.

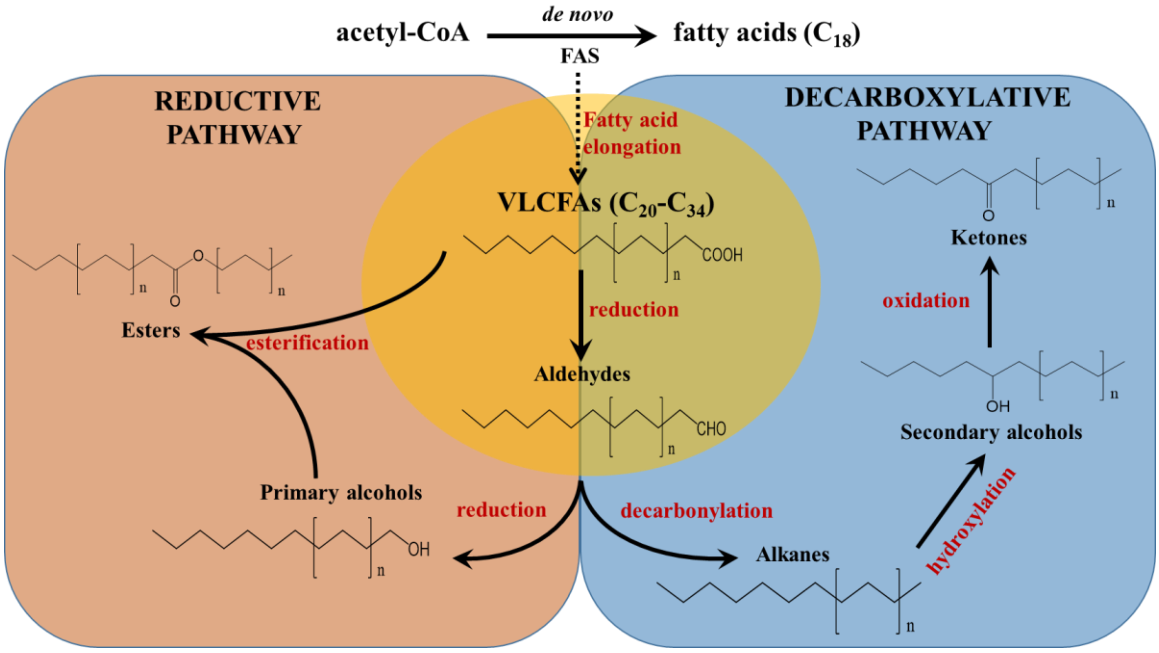


Figure 3. Proposed mechanism of epicuticular lipid biogenesis (Bianchi et al., 1985; Post-Beittenmiller, 1996). Very long chain fatty acids (>C20) produced by the FAE complex is reduced to fatty aldehydes which is a common intermediate of two pathways; reductive and decarboxylative pathways. In the decarboxylative pathway, the common aldehyde converts to alkanes, secondary alcohols and ketones via decarbonylation, hydroxylation and oxidation reactions respectively. The reductive pathway fatty aldehydes are reduced to primary alcohols which can be esterified with free fatty acids forming esters.

The biogenesis of the extracellular epicuticular lipids has been investigated by identifying metabolite intermediates and forward molecular genetic approaches that use *eceriferum* (*cer*) mutants in Arabidopsis (Koornneef et al., 1989) and *glossy* (*gl*) mutants in maize (Schnable et al., 1994) and tomato (Leide et al., 2007; Vogg et al., 2004) and *wax crystal-sparse leaf (wsl)* mutants of rice (Wang et al., 2017; Vogg et al., 2004). Based on these studies a model has been proposed (Figure 3) in which fatty acid elongation processes feeds two distinct fatty acid modification pathways: 1) a reductive pathway that generates fatty aldehydes, alcohols and wax esters; and 2) a decarboxylative pathway that converts the common aldehyde intermediate to hydrocarbons,

ketones and secondary alcohols (Bianchi et al., 1985; Post-Beittenmiller, 1996). Interestingly in maize, both these pathways are differentially expressed among different organs, where in seedling leaves the reductive pathway predominates (Bianchi et al., 1985) and in silks the decarboxylative pathway dominates (Loneman et al., 2017; Perera et al., 2010).

Enzymes that catalyze the abovementioned proposed model are localized on the endoplasmic reticulum (ER). The mechanism by which these lipids are transported and deposited as epicuticular waxes remains elusive. However, several mechanisms have been proposed to explain lipid trafficking from ER to plasma membrane (PM), and from PM to the extracellular matrix (for review see Bernard and Joubès 2013). Briefly, the intracellular trafficking is hypothesized to occur by physical contact between the ER and PM, via oleosine-like vesicles budding off from ER, using secretary vesicles of the Golgi apparatus and by acyl carrier proteins. The lipid export from PM to the extracellular matrix however has been interpreted based on gene mutation studies, which revealed the involvement of ATP binding cassette (ABC) transporters (ABCG 11 and ABCG12; McFarlane et al., 2010; Pighin et al., 2004) and lipid transporter proteins (LTPG1; DeBono et al., 2009). However both these transporters contribute to only a small portion of lipid transport.

Glossy2 is an exemplary product of forward genetics identified based on its characteristic "water beading" mutant phenotype (Hayes and Brewbaker, 1928). The initial epicuticular lipid profiles of *gl2* (Bianchi, 1975) identified that this mutation affected the fatty acid elongation - reductive pathway, specifically between C30 and C32. Molecular cloning of *Glossy2* revealed that this gene is homologous to Arabidopsis *CER2* (Tacke et al., 1995; Xia et al., 1996); sharing ~60% amino acid sequence similarity. Both these proteins are archetypal of BAHD acyltransferase enzymes (D'Auria, 2006). BAHD enzymes catalyze the acylation of alcohols or amine groups to

form ester and amide groups that contribute to large number of specialized metabolites (D'Auria, 2006; St-Pierre and Luca, 2000) (Figure 4).

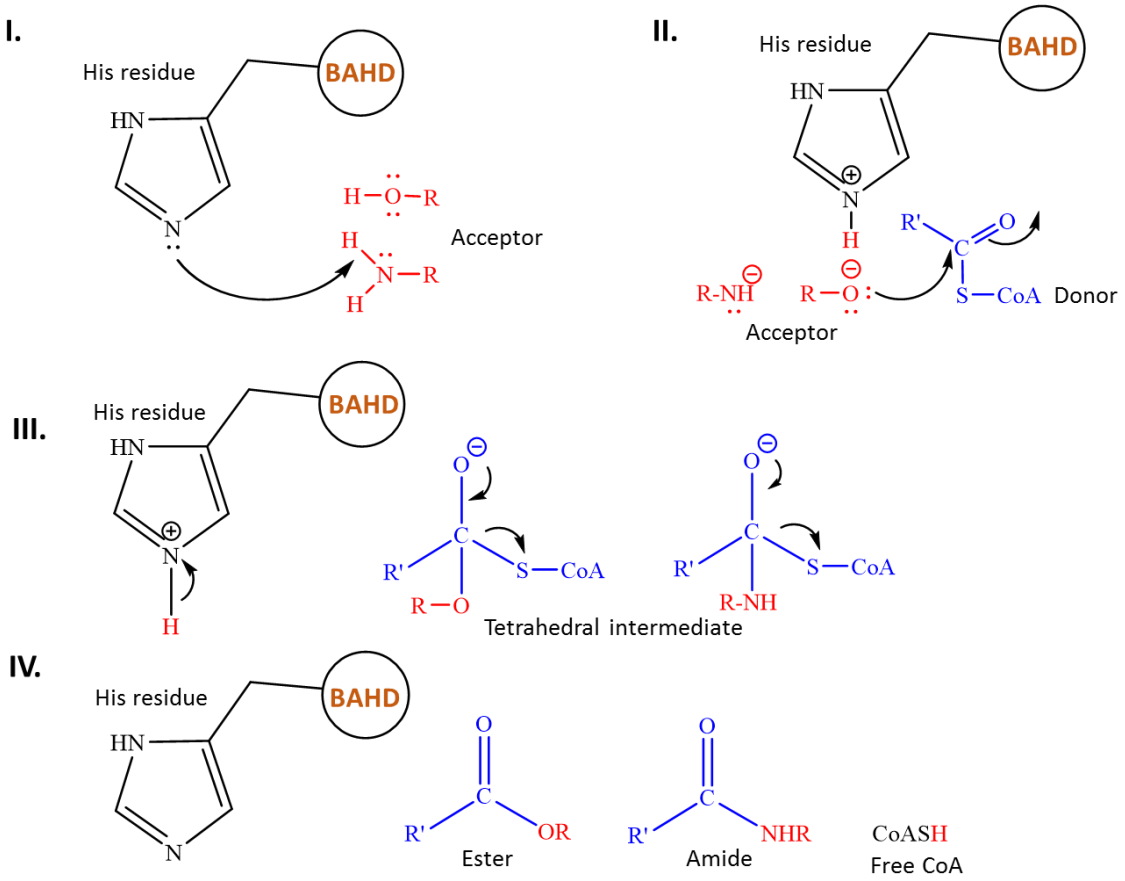


Figure 4. Schematic representation of BAHD enzyme catalyzed reactions. The histidine residue at the HXXXDX motif deprotonates the acyl acceptor (Step I), the resultant electronegative acyl acceptor carries out nucleophilic attack on carbonyl carbon of acyl CoA (donor) (Step II) forming a tetrahedral intermediate (Step III), and finally the release of metabolites containing ester or amide bonds between acceptor and donor and free CoA (Step IV).

The specific biochemical and physiological functions of the GLOSSY2/CER2 BAHD enzymes remains uncharacterized. However, recent yeast expression studies suggest that *CER2* (Haslam et al., 2012) and its rice homolog, *OsCER2* (Wang et al., 2017) associates with the fatty acid elongase (FAE) system. However, the mechanism by which such an interaction occurs and the role of BAHD catalytic motif remains unclear. Moreover, majority of the past genetic based studies on epicuticular lipids accumulation are limited to a single time-point of analysis, and

overlook the dynamics of the processes that integrate the developmental program of the tissues/organs that are being evaluated (Jenks et al., 1995; Goodwin et al., 2005; Pascal et al., 2013). Therefore, to address these complexities in epicuticular lipid biogenesis, this study employed the use of genetic manipulations via transgenic strategies and analytical mass spectrometry technologies (GC-MS, PTV-GC-MS, LC-qTOF and MALDI-MSI) to evaluate the functional ability of *Glossy2* and *Glossy2-like* on epicuticular lipids and the overall lipid metabolic network.

Dissertation organization

This dissertation is organized into 8 chapters and 2 appendix chapters.

Chapter 1 is a generalized introduction that outlines the concept of this dissertation as it pertains to mass spectral imaging in plant metabolism and epicuticular surface lipid biosynthesis.

Chapter 2 to 4 are published papers that are a joint collaborative effort between the Dr. Basil J. Nikolau group and the mass spectrometry imaging groups of Dr. Young-Jin Lee (Chapter 2 and 3) and Dr. R.S. Houk group (Chapter 4) at Iowa State University. These chapters employ analytical imaging technologies to answer biological questions. Specifically, **Chapter 2** and **Chapter 3** uses MALDI to spatially profile metabolic distributions during maize seed germination, and to study the distribution of thylakoid membrane lipids among individual cells of maize leaf, respectively. The mass spectral imaging was performed by Dr. Adam D. Feenstra and Dr. Maria E. Duenas. The growing of biological material, microscopy analysis, and metabolic GC-MS analysis was performed by Liza E. Alexander and partial analysis of GC-MS was also performed by Dr. Zhihong Song. **Chapter 2** is a shared first-author paper between Adam D. Feenstra and Liza E. Alexander. **Chapter 4** uses ELDI to study the spatial distribution of metabolites in pigmented coleus plants. The spatial imaging was performed by Dr. Patrick A. McVey. Plant

growth, experimental design, and microscopy analysis was performed by Liza E. Alexander. Metabolic profiling was performed via LC-MS and GC-MS. The LC-MS analysis was performed by Dr. Bo Xie, and the GC-MS analysis was performed by Xinyu Fu and Liza E. Alexander. **Chapter 4** is a shared first-author paper between Patrick A. McVey and Liza E. Alexander.

Chapter 5 to 7 are manuscripts in preparation. They focus on the functional characterization of specific maize genes, *Glossy2* and *Glossy2-like*, and their involvement in epicuticular lipid biosynthesis. Specifically, **Chapter 5** is the *in planta* characterization of *Glossy2* homologs using a transgenic approach in Arabidopsis. All experiments were performed and analyzed by Liza E. Alexander with the assistance of undergraduate assistants – Aerial Davis, and Michael A. Schelling. A section of this chapter pertaining to lipidomics was performed at Dr. Kazuki Saito's lab, RIKEN, Japan under the guidance of Dr. Yozo Okazaki. This was a part of the NSF-EAPSI grant awarded to Liza E. Alexander. **Chapter 6** is the mass spectral imaging of epicuticular lipids in developing Arabidopsis flowers that carry the *Glossy2* and *Glossy2-like* transgenes. Plant growth and single flower epicuticular lipid analysis was performed by Liza E. Alexander. Mass spectral imaging of individual flowers was performed by Dr. Bo Xie and undergraduate assistant Jena S. Gilbertson. **Chapter 7** is the yeast heterologous expression of the *Glossy2* and *Glossy2-like* genes to evaluate their roles in fatty acid elongation. A few of the strains were provided or developed by graduate colleagues Kayla Flyckt and Kenna E. Stenback. Growth and metabolic GC-analysis of yeast fatty acid profiles were performed by Liza E. Alexander with the assistance of undergraduate students- Michael A. Schelling and Katelyn R. Campbell.

Chapter 8 discusses a general conclusion of this body of work and the future prospects of this work.

Appendix A includes the co-expression of KCS isozymes with *Gl2* homologs mutated at their BAHD catalytic domain in yeast, which is not included in **Chapter 7**.

Appendix B includes the purification of GLOSSY2 protein via heterologous expression in *E.coli*.

CHAPTER 2. SPATIAL MAPPING AND PROFILING OF METABOLITE DISTRIBUTIONS DURING GERMINATION OF MAIZE SEEDS

[1]Modified from paper published in Plant Physiology (2017) 174 (4): 2532-2548

Adam D. Feenstra[a,b,#] Liza E. Alexander[b,c,d,#], Zhihong Song[d], Andrew R. Korte[a,b], Marna D. Yandeau-Nelson[d,e], Basil J. Nikolau[b,c,d], and Young Jin Lee[a,b]

[#]These authors contributed equally to the work

Abstract

Germination is a highly complex process by which seeds begin to develop and establish themselves as viable organisms. In this study, we utilize a combination of gas chromatography-mass spectrometry, liquid chromatography-fluorescence, and mass spectrometry imaging approaches to profile and visualize the metabolic distributions of germinating seeds from two different inbreds of maize (*Zea mays*) seeds, B73 and Mo17. Gas chromatography and liquid chromatography analyses demonstrate that the two inbreds are highly differentiated in their metabolite profiles throughout the course of germination, especially with regard to amino acids, sugar alcohols, and small organic acids. Crude dissection of the seed followed by gas chromatography-mass spectrometry analysis of polar metabolites also revealed that many compounds were highly sequestered among the various seed tissue types. To further localize

[1] [a]Department of Chemistry, Iowa State University, Ames, USA
[b]Ames Laboratory-US DOE, Ames, IA, USA
[c]Roy J. Carver Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, IA, USA
[d]Center for Metabolic Biology, Iowa State University, Ames, IA, USA
[e]Department of Genetics, Development & Cell Biology, Iowa State University, Ames, IA, USA

compounds, matrix-assisted laser desorption/ionization mass spectrometry imaging was utilized to visualize compounds in fine detail in their native environments over the course of germination. Most notably, the fatty acyl chain-dependent differential localization of phospholipids and triacylglycerols was observed within the embryo and radicle, showing correlation with the heterogeneous distribution of fatty acids. Other interesting observations include unusual localization of ceramides on the endosperm/scutellum boundary and subcellular localization of ferulate in the aleurone.

Introduction

Plants use seeds as the propagule to ensure reproduction to the next generation, and over the past 10,000 years human civilizations have established agricultural practices to ensure a seed-based food supply (Larson et al., 2014). Therefore, deciphering the processes that enable seeds to perform their biological functions is of importance in understanding how plants are propagated, and is also of practical importance to improve agriculture. Seeds are designed to survive long periods of dormancy in a relatively dry state, and the process of germination is initiated by the imbibition of water. During this germination process, many metabolic changes occur, most of which are associated with the catabolism of seed storage products (proteins, polysaccharides, and lipids) into metabolically useable, simpler chemical forms that are either used as precursors to assemble the growing seedling or are oxidized via energy-producing biochemical pathways to thermodynamically support growth (Bewley, 1997, 2001).

The specific metabolic processes that support seed germination are somewhat dependent on the taxonomic clade of the plant, which determines seed tissue/organ organization and the nature of the seed storage compounds. Specifically, the seeds of the Poaceae family of monocots are characterized by a starch- and protein-filled endosperm, and their catabolism by digestive

enzymes produced in the outermost layer of the endosperm (i.e. the aleurone) provides the carbon-based and nitrogenous precursors for the growth of the embryo, which is comprised of the embryonic axis and the scutellum (Dante et al., 2015). Thus, during seed germination the glycolytic and oxidative pentose phosphate pathways are induced to metabolize the hydrolyzed starch. Smaller quantities of seed storage proteins are also catabolized to provide the amino acid precursors for new protein synthesis in the embryo, or carbon skeletons for reassimilation into anabolic processes, or further catabolized for energy generation. Finally, lipid in the form of triacylglycerol (TAG) is stored in a specialized tissue (in maize [*Zea* mays], the scutellum), and its catabolism via the β-oxidation of fatty acids can provide the elongating embryo axis an energy source. Alternatively, coupling with the glyoxylate pathway, which is induced during germination, carbon from fatty acids can be used to assemble new metabolic intermediates to support embryo growth (Firenzuoli et al., 1968; He et al., 2015).

Thus, seed germination requires the coordinated induction of a number of processes that are nonuniformly distributed among the tissues and organs of the seed. Many of the metabolic pathways that are induced during the seed germination processes that involve starch, oil, and protein turnover have been studied via molecular, genetic, and biochemical studies (Ingle et al., 1964; Limami et al., 2002; Shu et al., 2008; Zhang et al., 2009). Therefore, techniques that allow for the determination of spatial localization of metabolites in fine detail, specifically mass spectrometry imaging (MSI), should allow for further insight into the mechanisms and actions of the pathways that are integral to seed germination.

MSI has become an increasingly powerful technique with which to study the spatial distribution of biological molecules within tissues (McDonnell and Heeren, 2007). In a typical MSI experiment, a sampling probe is rastered across a tissue sample, with mass spectra collected

at each raster point. Following data acquisition, the spatial distribution of ions of interest can be visualized based on the x-y coordinates of the collected spectra. Various sampling probes have been adopted for MSI along with accompanying ionization methods (McDonnell and Heeren, 2007; van Hove et al., 2010). Owing to its ease of use, high sensitivity, and ability to ionize a wide variety of molecules, matrix-assisted laser desorption/ionization (MALDI) has been the most widely used MSI technique. MALDI-MSI can obtain high-spatial-resolution images of ions, allowing for visualization of fine structures. Spatial resolution of ~20 μm size has become routine in MALDI-MSI, and 2- to 5-μm spatial resolution has been demonstrated by several groups, with the latter providing a means for locate metabolites within subcellular compartments (Zavalin et al., 2013; Korte et al., 2015).

Recently, MALDI-MSI has begun to see increased application in the study of plant biology (for review, see Kaspar et al., 2011; Lee et al., 2012). Here, we applied MALDI-MSI in combination with gas chromatography-mass spectrometry (GC-MS) and HPLC-based analyses of extracts to study the distributions and profiles of metabolites in germinating maize seeds at four different time points after imbibition. Two contrasting maize genotypes (inbreds B73 and Mo17) were compared during germination to explore the metabolic differences arising from the genetic differences between these two inbreds. GC-MS and HPLC-based metabolite profiling of extracts prepared from the physically dissected seeds was used to distinguish between the metabolomes of the two inbreds during early germination and provide a quantitative validation of the MSI localization data. Three different MALDI matrices were used in either positive or negative ion mode to image the spatial distribution of a wide range of metabolites. This integrated approach provides unprecedented insights into the spatially resolved coordination of metabolic processes that are sequestered among the germinating embryo axis, the lipid-rich scutellum, the nutritive

endosperm, the digestive aleurone, and outer pericarp cell layers during the early stages of seed germination.

Results

Overview of Experimental Workflow

A schematic of the experimental workflow is provided in Figure 1. Maize seeds were collected from the B73 and Mo17 inbreds at four time points during the early phases of germination (0.2, 12, 24, and 36 h postimbibition). At each time point, nine seeds comprising three biological replicates (three seeds each) were immediately flash frozen and extracted for whole-seed metabolite profiling. These extracts were analyzed via a GC-MS nontargeted global metabolite analysis platform and via a liquid chromatography (LC)-fluorescence-based targeted analysis platform for amino acids. In addition, at the 12- and 36-h time points, nine seeds comprising three biological replicates (three seeds each) were collected for GC-MS analysis of seeds dissected into component endosperm, embryo, tip cap, and pericarp. In parallel, three additional seeds from each of the four time points were flash frozen in liquid nitrogen and subjected to MSI analysis. These seeds were sectioned with 10 μm thickness (25 sections per seed), and four sections (5, 11, 17, and 22) were inspected via optical microscopy. The three sections nearest to the most intact section as determined via optical microscopic inspection were used for MSI analysis. Microscopic analysis revealed detailed morphological features of the seed, including the pericarp, which is the outermost layer of the seed, the aleurone layer of the endosperm, the endosperm itself, the embryonic axis, the scutellum, and the tip cap. As germination progressed, the radicle also became visible as it protruded from the embryonic axis of the embryo. It is important to note that it can be difficult to distinguish between the multicellular pericarp and the single-cell aleurone layer at lower

magnification. In the cases where these cell layers cannot be distinguished, we will refer to them as the aleurone and/or pericarp.

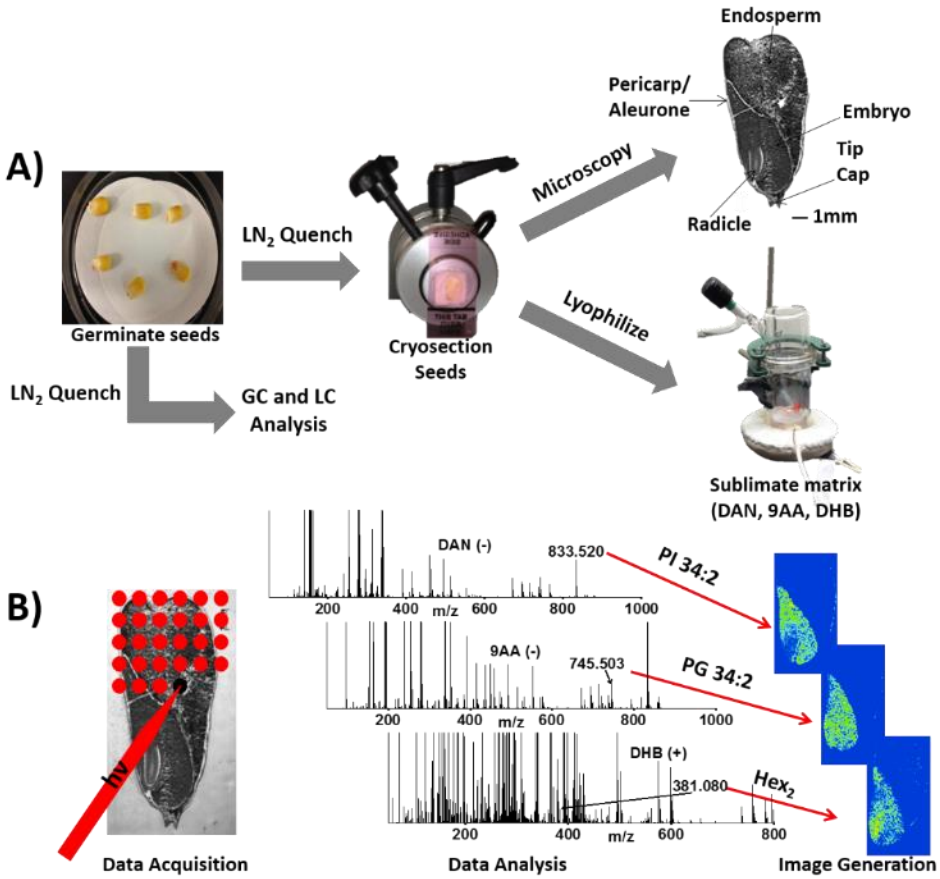


Figure 1 Illustration of experimental workflow. **(A)** Sample preparation: Seeds were germinated in a petri dish and quenched at the desired time point with liquid nitrogen (LN_2). Some samples were selected for metabolite profiling analysis. The remaining samples were cryo-sectioned at 10 µm thickness for microscopic and MSI analysis. **(B)** MSI data acquisition and analysis: The laser was rastered across the tissue sample, collecting a spectrum at every x-y position. Resulting mass spectra were evaluated and MS images were generated for individual ions at selected *m/z* values.

The analytes that can be characterized through MSI often are significantly limited by the choice of matrix and ion polarity. In this work, MSI analysis was performed on the three tissue sections with one of three different matrices, 1,5-diaminonaphthalene (DAN), 9-aminoacridine (9AA), and 2,5-dihyroxybenzoic acid (DHB), to cover a relatively wide range of analytes. DAN and 9AA were used in negative ion mode (mass-to-charge ratio [*m/z*] range 50–1,000), while DHB was used in positive ion mode (*m/z* range 50–800 and *m/z* range 600–1,600). In total, 72

seed sections were analyzed by MSI (2 inbred lines × 3 seeds per time point × 4 time points × 3 sections per seed for each of the three matrices). A spatial resolution of 100μm was used to minimize the total data acquisition time in this large-scale experiment while still providing sufficient spatial resolution to match the morphological features observed in the optical images.

Over the course of the analyses, hundreds of analytes were observed by both MSI and the gas chromatography (GC)- and LC-based metabolite profiling platforms. However, due to biological and analytical variations, many low-abundance analytes were not reproducibly detected across all data sets. Therefore, we limit our discussion to the more abundant analytes that were consistently observed across three biological replicates. For simplicity of presentation of the MSI data, all the *m/z* values and mass tolerances used in producing images are summarized in Supplemental Table S1, as are the color scales used to produce false color images.

Analysis of the metabolomes of germinating seeds

The nontargeted GC-MS metabolomics analyses of the whole seeds at four germination time points detected 162 analytes, of which 63 were chemically identified. These metabolites include sugars (monosaccharides and disaccharides), sugar acids, organic acids, phenolics, nitrogenous metabolites, polyols, esters, lipids, fatty acids, and sterols (Supplemental Table S2). There are clear differences in the metabolome between the two inbreds, and the metabolite profile also is affected by the germination process. Using log-ratio plots, we evaluated the degree to which these differential metabolites between the inbreds are differentially affected by the germination process (Fig. 2A). In these comparisons, we used the 12-h time point as the anchor for all comparisons among genotypes and the seed germination time line. At this anchor time point, there are 63 metabolites that accumulate at significantly different levels between the inbreds. By comparing such log-ratio data at each of the four time points, one can gain insights into how the

germination process affects the metabolome. One readily recognizable class of metabolites that are clearly distinct between the two inbreds is the amino acids, and this platform detected 10 of the proteinogenic amino acids. Therefore, we confirmed these differences with a second analytical platform, which specifically targeted amine-containing metabolites with LC-fluorescence, and this detected 18 of the proteinogenic amino acids (Supplemental Table S3), the exceptions being Pro and Cys. In both the targeted LC-amino acid platform (log-ratio plot in Supplemental Fig. S1) and the nontargeted GC-MS metabolomics platform (Fig. 2A), most amino acids are at significantly higher levels in Mo17 seeds than in B73 seeds. These data are consistent with previous studies of these inbreds that identified Mo17 mature seeds having a higher amino acid content than B73 seeds (Römisch-Margl et al., 2010). As the germination process proceeds, these differences in amino acid content between the two inbreds is reduced, so that by 36 h postimbibition, the profiles are almost identical except for Lys, Thr, and Arg, which are still slightly higher in Mo17, and Trp, which is now slightly higher in B73 (Supplemental Fig. S1).

More broadly, 64% of the detected metabolome (i.e. 104 metabolites) accumulates to different levels between the two inbreds at least at one time point during the germination process ($P < 0.05$ in Supplemental Table S2). Among the 104 metabolites that accumulate differentially among any of the time points, 21 are differential at all time points evaluated (Fig. 2B) and 15 are differential among any of three consecutive time points. Seven of the first group are chemically identified and appear to be intermediates of metabolic processes that would be expected to be hyperactive during seed germination. These are specifically associated with carbohydrate metabolism (i.e. sorbitol, cellobiose, rhamnose, and talose), which may be associated with starch mobilization that is occurring during the breakdown of the endosperm tissue or possibly cell wall deposition as the new seedling tissues begin to be assembled. Eight of the chemically identified

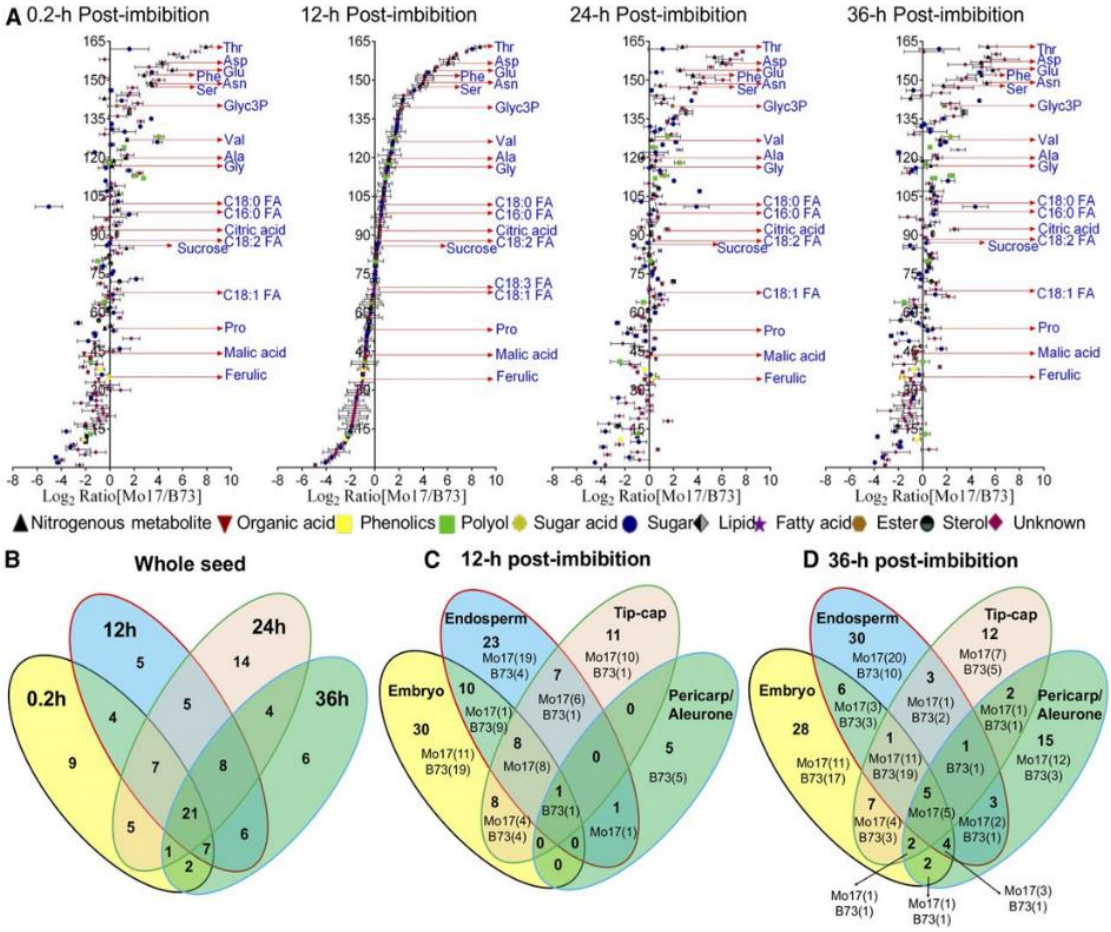


Figure 2 Log-ratio plot comparison **(A)** and Venn-diagram representations of the differential metabolomes between Mo17 and B73 maize inbred of whole seeds **(B)** and micro-dissected organs from seeds at 12-h **(C)** and 36-h **(D)** post-imbibition. In the log-ratio plots, the X-axis plots the log-transformed relative abundance ratio of each metabolite in Mo17 versus B73. The Y-axis plots the individual metabolites (162 analytes, 63 chemically defined), and the order of the metabolites on the Y-axis is identical and ordered from the lowest to the highest value on the X-axis as determined for the 12-h post-imbibition time point. The arrows identify amino acids or those metabolites also analyzed by MSI. Glyc3P = glycerol 3-phosphate; FA = Fatty acid. The Venn-diagram in B represents the distribution of metabolites that are differentially expressed between Mo17 and B73 seeds ($p<0.05$) among the four post-imbibition time points. The Venn diagrams in C and D show metabolites that are differentially expressed between Mo17 and B73 seeds ($p<0.05$) in micro-dissected organs from germinating seeds at 12-h and 36-h post-imbibition. The identity of the inbred and the number of the metabolites that occur at higher levels is represented by the labels inbred(x).

metabolites that are differential among the three consecutive time points are amino acids (i.e. Asp, Glu, and Thr) in the three early time points and lipids (i.e. stigmasterol, sitosterol, and linoleic acid) that are differential among the three latter time points. These profiles are consistent with seed storage protein mobilization, which occurs at the earlier stages of seed germination, and membrane deposition, which occurs later in the process as the seed radicle begins to grow and emerge. Finally,

among these chemically identified differentially expressed metabolites are succinic acid and fumaric acid, which are intermediates of the tricarboxylic acid cycle, and these are associated with the high rates of respiration that are needed to support the germination process (Bewley et al., 2013). Collectively, these data indicate that these metabolic processes are differential between the two inbreds, which correlates with the phenotypic observation that these two inbreds germinate at different rates.

Bearing in mind that a germinating maize seed consists of different tissues, each of which expresses diverse metabolic processes, we evaluated the distribution of the metabolome among four distinct seed organs that are readily separable by microdissection (i.e. the pericarp and aleurone layers, the embryo, the endosperm, and the tip cap). The metabolomes of these microdissected organs were evaluated at 12 and 36 h postimbibition. Because of the wider range of metabolites that were detected, the GC-MS analysis of the metabolomes of the microdissected seed organs focused on the polar class of metabolites. These analyses revealed the relative abundances of 218 analytes, of which 75 were chemically defined (Supplemental Table S4). Supplemental Figure S2 shows the log-ratio plot for the relative abundances of all analytes between Mo17 and B73, comparing the metabolomes of each dissected organ and the metabolomes of the whole seed.

An obvious advantage in the microdissected data set is that it reaches lower abundance metabolites, revealing the relative abundances of an extra 56 analytes (Supplemental Table S4) that were not detected when whole seeds were analyzed. More significantly, the metabolites that are the most differential between the metabolomes of the whole seed are different when one considers the metabolomes of the separated organs, which is indicative of the different metabolic processes that are being expressed in these individual organs. At both time points, the endosperm

and embryo tissues account for the majority of the differentially expressed metabolites. About half of these are shared among multiple evaluated organs, but about 20% of the metabolome is uniquely differential in either the embryo or endosperm tissues (Fig. 2, C and D). The endosperm-specific differential metabolites are primarily amino acids at 12 h postimbibition, but by 36 h, these are a mixture of amino acids, sugars, and organic acids. This is consistent with the degradation of the starch and seed storage proteins that are concentrated in the endosperm of maize seeds and indicates different rates of their catabolism between the two inbreds.

Therefore, these data indicate that, as the seed germination process proceeds, the expressed metabolome changes between inbred lines, with larger changes occurring in the embryo, followed by the endosperm and tip cap. Fewer differentially expressed metabolites occur in the metabolome of the pericarp and aleurone tissues. More refined analyses of the nonuniform distribution of the metabolites in the different tissues of the germinating maize seed were obtained by MSI.

Heterogeneous distribution of the mobilization of seed storage reserves

In the grasses, energy in dormant seeds is stored predominantly as polysaccharide starch granules within the endosperm, which can be mobilized by hydrolytic enzymes (e.g. α- and β-amylases) released by the aleurone and scutellum of germinating seeds (Zeeman et al., 2007). The images in Figure 3A show large hexose polysaccharides (Hex_5–Hex_9), presumably degraded from starch, observed at low abundance levels and localized in the endosperm. Disaccharides, predominantly Suc, also are observed and are present at much higher levels and primarily constrained to the germinating embryo. At the later stages of germination, the disaccharide signal appears to be concentrated more extensively in the emerging radicle of the embryonic axis. Disaccharide also is detected within other tissues, such as the endosperm and pericarp and/or aleurone, but at lower abundance. These nonhomogenous distributions of the polysaccharides are

conserved in both B73 and Mo17 inbreds and are consistent with the mobilization of Suc within the scutellum and transport to the embryonic axis, specifically the radicle (Sánchez-Linares et al., 2012).

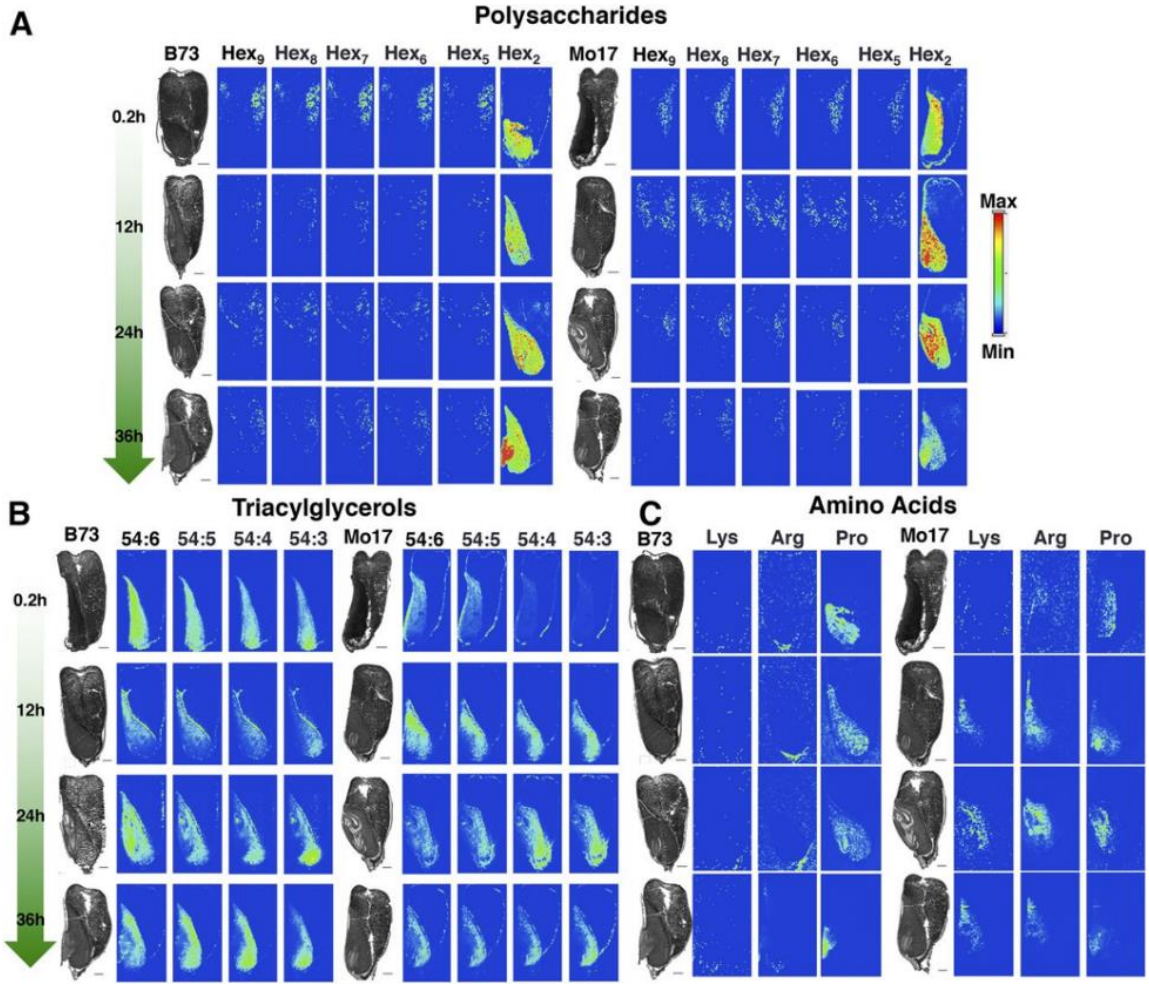


Figure 3 MSI images of hexose polysaccharides (A), TAGs (B), and amino acids (C) in germinating seeds. All ions were detected in positive ion mode with DHB as the matrix. Large polysaccharides were detected as potassium adducts of water loss, $[M\text{-}H_2O+K]^+$, likely due to in-source fragmentation during MALDI-MS data acquisition. Disaccharide and TAGs were detected as potassium adducts, $[M+K]^+$, and amino acids were detected as protonated ions, $[M+H]^+$. Bar = 1 mm.

Seed oil energy reserves in maize are localized to the embryo, which is in contrast to other cereal grains (e.g. wheat [*Triticum aestivum*] and oat [*Avena sativa*]) that store oil in the endosperm (Leonova et al., 2010). The embryonic localization of four TAG species differing in fatty acyl chain composition is shown in Figure 3B. At the initial stages of germination (i.e. 0.2 and 12

h), TAGs are distributed homogenously throughout the embryo, including the embryonic axis and the scutellum. As germination progresses and the radicle of the embryonic axis begins to elongate, TAGs begin to display nonuniform localization patterns, similar to fatty acid and phospholipid molecular species, as will be discussed later.

During the germination process, amino acids become available from the breakdown of storage proteins initiated by the hydrolytic activity of protease enzymes synthesized in the protein storage vacuoles. The reduced nitrogen associated with these amino acids can be used for additional de novo synthesis of amino acids (Bewley et al., 2012). Because of their low ionization efficiency and amphiprotic nature, amino acids are difficult to analyze by MALDI-MSI (Toue et al., 2014). Despite this known difficulty, three amino acids were reliably detected in the MSI experiments, these being Lys, Arg, and Pro (Fig. 3C). Lys and Arg are barely detectable in B73 seeds but are clearly visible in Mo17 seeds. This striking difference in the abundance of Lys and Arg between the two inbreds is consistent with the parallel profiling data of these amino acids generated by LC-fluorescence analysis of *o*-phthaldialdehyde (OPA) derivatives on extracts of seeds (Supplemental Fig. S1); the concentrations of these amino acids are significantly lower in extracts of B73 seeds than Mo17 seeds throughout the germination process. The LC method could not analyze Pro levels in these extracts because the OPA derivatization requires a primary amine group. Rather, Pro concentrations were compared from parallel GC-MS analyses of the polar extracts of seeds (Fig. 2A; Supplemental Fig. S2). These data are in agreement with the MALDI-MSI data, with Pro levels being similar or slightly higher in B73 seeds than in Mo17 seeds. The MSI data establish that the accumulation of these amino acids (Lys, Arg, and Pro) is concentrated in the embryonic axis and the scutellum of the seed embryo. Moreover, at later time points, Pro is more prominent in the radicle of the seed, but Lys and Arg appear to be less abundant in the radicle.

These MSI-based visualizations of the in situ localization of these amino acids are consistent with the direct quantitative measurement of the levels of these amino acids in extracts of microdissected embryo tissues of the maize seeds (Supplemental Fig. S3).

Metabolites enriched in pericarp/aleurone layers

The pericarp and aleurone layers of the seed are responsible for the protection of the mature seed as well as the initial uptake of water that begins germination. As germination begins in maize, many of the enzymes needed to break down storage molecules in the seed are released from the aleurone layer of the endosperm (Chrispeels and Varner, 1967; Fincher, 1989; Bernier and Ballance, 1993). In this work, three metabolites are found to be uniquely located at the perimeter of the cross section and, therefore, can be surmised to be located either in the pericarp (the outermost layer of the kernel) or the aleurone layer, which is the outermost single cell layer of the endosperm. The three metabolites remained localized to this region of the seed throughout the germination process and were not detected in the embryo or the endosperm (Fig. 4A). This is in contrast to most other metabolites, such as Suc (Fig. 3A) or citrate (Fig. 5), that are found in the aleurone and/or pericarp but also in other tissues. One of these metabolites (*m/z* 193.051) was identified as ferulate by accurate mass (less than 5 ppm) and tandem mass spectrometry (MS/MS; Supplemental Fig. S4). However, useful MS/MS spectra could not be obtained for the other two ions (*m/z* 135.046 and 161.025) because of their low abundance. The chemical compositions calculated from the accurate masses, $C_8H_8O_2$ and $C_9H_6O_3$, respectively, suggest that they may have at least one aromatic ring.

The spatial resolution used in this experiment, 100 μm, was not sufficient to determine whether these metabolites are localized specifically to the multiple cell layers of the pericarp or the single cell layer of the aleurone. To further explore the exact location of these compounds, a

high-resolution MSI experiment was performed with 10 µm spatial resolution on a small region of the outer perimeter of B73 maize seeds at 24 h postimbibition (Fig. 4, B and C). Because of the limited sensitivity inherent to a much smaller sampling area, only ferulate provided sufficient ion signal to produce a clear image. The ferulate image was compared with that of the phospholipid phosphatidylinositol (PI) 34:2 and malate to better understand the relative localization of these compounds. Comparing these molecular images with the corresponding optical image (Fig. 4B) and overlaying MS images with the optical image (Fig. 4C), ferulate (red) appears to be present exclusively in the aleurone layer, particularly at the boundary between the aluerone and the pericarp, while the phospholipid PI 34:2 (green) is present throughout the aleurone layer and malate (blue) is localized in the endosperm and the pericarp layer of the seeds.

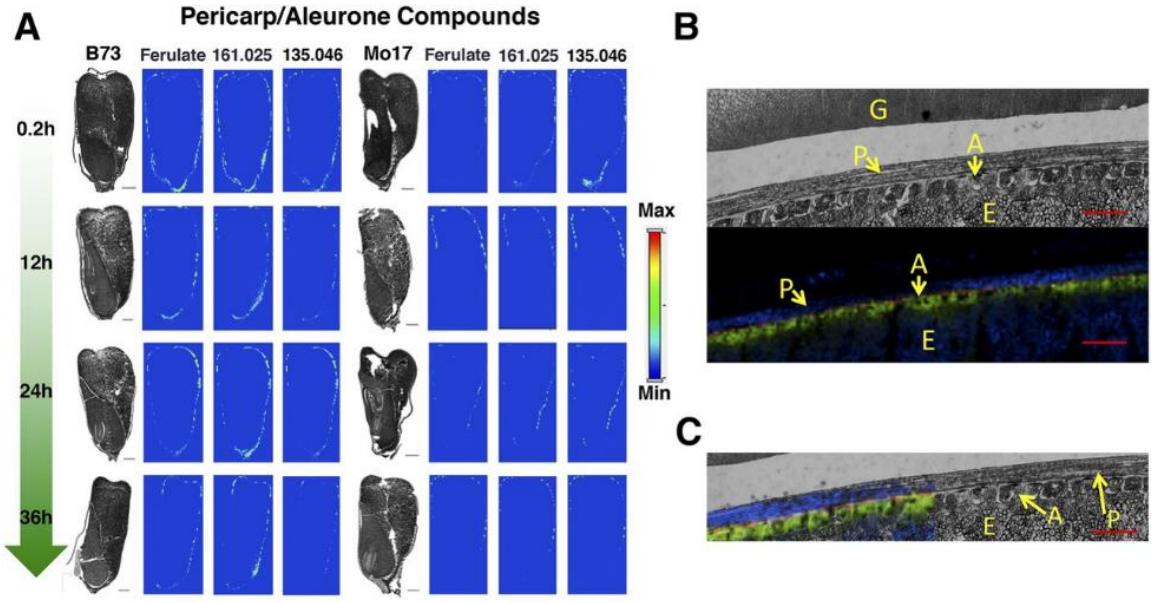


Figure 4 (A) MSI of three compounds uniquely localized to the perimeter of the seed. Images were acquired in negative ion mode with DAN as the matrix. No normalization was applied to these images. Bar = 1 mm. **(B)** Optical microscopic (top) and MS (bottom) images of the pericarp and aleurone layers of a B73 maize seed with 24-h post-imbibition. Morphological features are labeled on the images. A, aleurone.; E, endosperm; G, gelatin embedding medium; P, pericarp.. **(C)** Overlay of optical and MS images in B. MS images were obtained with DAN as the matrix in negative mode. MS images in B and C are ferulate (red), malate (blue), and PI 34:2 (green). All analytes were detected as deprotonated species, $[M-H]^-$. Bar = 100 µm.

Localizations of Respiratory Intermediates: Phosphorylated Metabolites and Organic Acids

During the process of seed germination, the respiration of seed reserves enables energy generation in the absence of photosynthesis, supporting the development of the seedling until it protrudes from the soil and is able to harness solar energy via photosynthesis. Respiration initiates within minutes of imbibition by the activation of cytosolic enzymes for glycolysis, the oxidative pentose phosphate pathway, and enzymes for the tricarboxylic acid cycle in mitochondria (Bewley et al., 2013). The distributions of two phosphorylated intermediates of respiration, hexose phosphate and glycerol phosphate, and one or possibly two organic acids of the tricarboxylic acid cycle, citrate and/or isocitrate, are shown in Figure 5. The phosphorylated metabolites are most likely Glc-6-P and glycerol-3-phosphate, respectively, but we cannot distinguish among the possible regioisomers or stereoisomers by MALDI-MSI or MS/MS experiments. Similarly, we are also unable to distinguish between citrate and isocitrate, and the resulting image is likely an integration of the distribution of these two metabolites.

In both inbreds, hexose phosphate and glycerol phosphate are localized in the germinating embryo and the aleurone and/or pericarp of the seeds, and this location is unaltered throughout the germination process. Hexose phosphate appears to be slightly less abundant in the radicle, whereas glycerol phosphate is more homogenously distributed throughout the embryo. The organic acids, citrate/isocitrate, are observed first at 0.2 h postimbibition, and they are concentrated predominantly in the perimeter (i.e. the aluerone and/or pericarp) of the seed. As germination progresses, citrate/isocitrate also occurs within the radicles and scutella of both inbreds, similar to the hexose phosphate and glycerol phosphate compounds; the former is probably associated with the induction of the glyoxylate pathway or tricarboxylic acid cycle, which is induced during

germination to convert lipid-derived carbon to sugars via an acetyl-CoA intermediate, while the latter two are an intermediate and side product of glycolysis, respectively (Bewley et al., 2013).

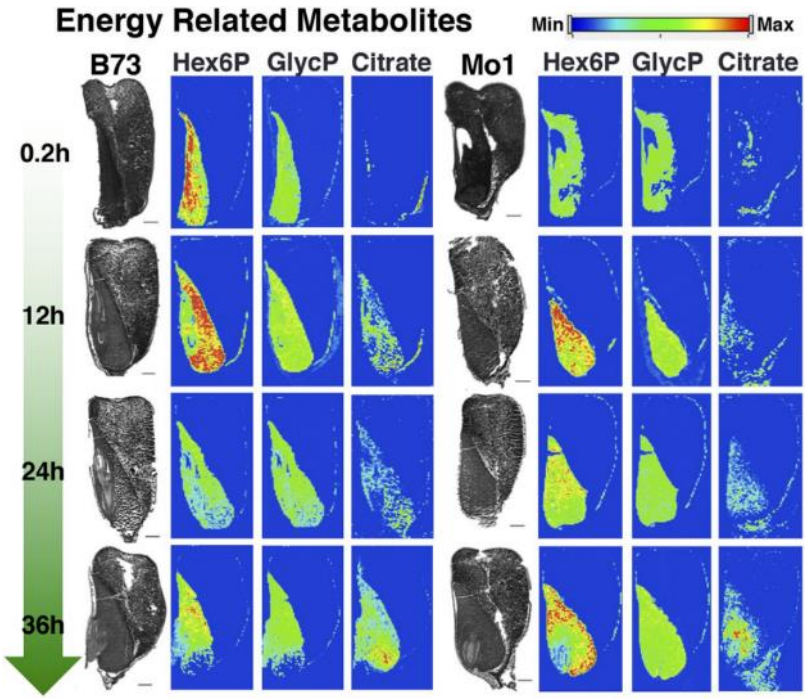


Figure 5 Distribution of hexose phosphate (Hex6P), glycerol phosphate (GlycP), and citrate/isocitrate. Phosphorylated metabolites are detected as deprotonated water-loss species, $[M\text{-}H_2O\text{-}H]^-$, in negative ion mode with DAN as the matrix. Citrate/isocitrate was detected as a deprotonated species, $[M\text{-}H]^-$, in negative ion mode with 9AA as matrix. Bar = 1 mm.

Malate is another molecule involved in the tricarboxylic acid cycle and is detected in all replicates at all time points. Unlike citrate, its localization is found only in the endosperm of the seed and radicle of some more fully developed seeds (Supplemental Fig. S5A). This is not compatible with GC-MS analysis of the microdissected seed samples, which indicated that most of the malic acid is present in the embryo of the seed rather than the endosperm (Supplemental Fig. S5C). In contrast, citrate is much more abundant in embryos (Fig. 5), matching well with the dissected GC-MS data (Supplemental Fig. S5D). According to our previous experience on MSI of leaf tissues, it is often difficult to obtain reproducible results with malate unless the tissues are prepared fresh and analyzed as soon as possible. Hence, we suspect that this observation may be

an artifact coming from the degradation of malate during tissue storage. To confirm this hypothesis, a separate MSI experiment was conducted on a seed that was analyzed on the same day it was collected. In this experiment, we found the majority of the malic acid in the embryo as expected (Supplemental Fig. S5B).

Differential distributions of lipid classes accentuated in the emerging radicle

In monocots such as maize, the embryonic axis develops into two main parts: the radicle, which is the primary root and is protected by the coleorhiza, and the plumule, or collection of leaf primordia, which is protected by the coleoptile. The germination process transitions into seedling growth with the protrusion of the radicle from the embryo. These anatomical structures develop with significant cell extension with or without cell division. Radicle growth through cell extension is a turgor-driven process that will involve the assembly of new membranes, requiring lipid biosynthesis (Bewley, 1997, 2001). Therefore, we applied MALDI-MSI to localize lipids during the germination process. MALDI-MSI efficiencies are dramatically different for each lipid class, depending on their chemical functionalities. Three matrices were used in positive and negative ion modes to visualize different classes of lipids. DHB was used in positive ion mode for phosphatidylcholine (PC) and ceramides; 9AA was used in negative ion mode for fatty acids; and DAN was used in negative ion mode for phosphatidylethanolamine (PE), phosphatidic acid (PA), and PI.

Figure 6 compares the spatial distribution of four fatty acids, palmitate (16:0), stearate (18:0), oleate (18:1), and linoleate (18:2), between the two inbreds at four germination stages. Over the course of the germination process, palmitate and linoleate occur homogenously throughout the embryo in both inbreds. In contrast, stearate is partially absent and oleate is almost completely absent from the radicles of both inbreds. This unprecedented finding is similar to the localization

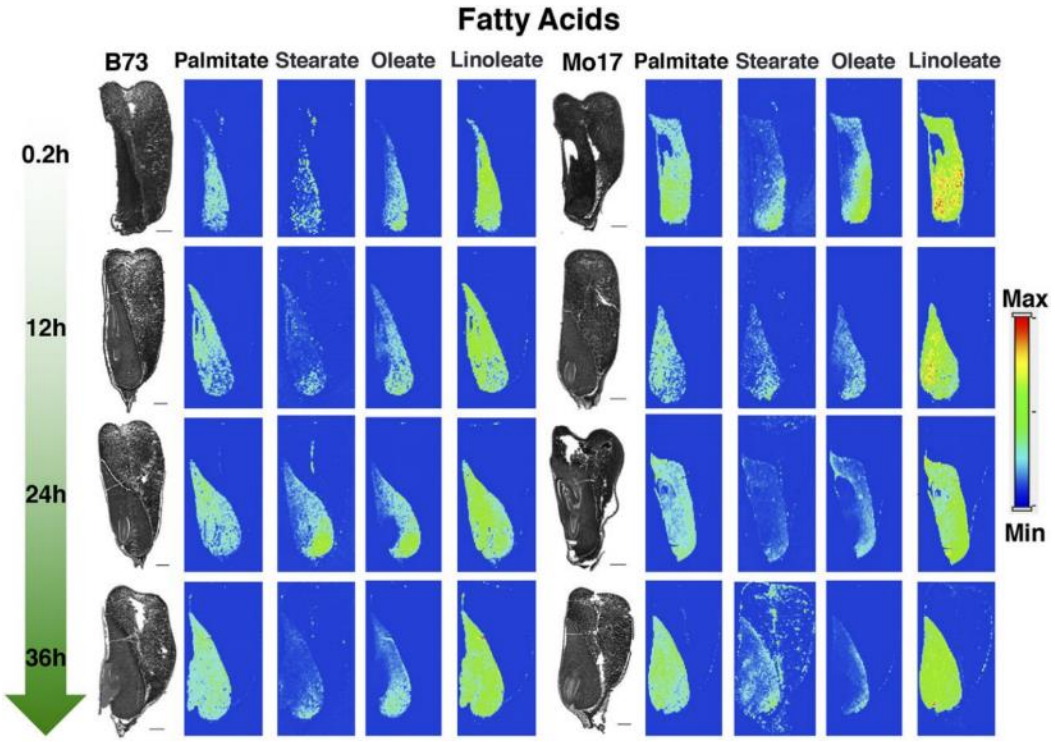


Figure 6 Distribution of fatty acid species. All ions were detected as the deprotonated species, $[M-H]^-$, in negative ion mode with 9AA as the matrix. Stearate has some contamination from vacuum pump oil. Bar = 1 mm

of TAGs (Fig. 3) and phospholipid classes (Fig. 7). However, it is in contrast to the GC-MS metabolomics data, which do not provide any spatial localization information for these fatty acids between the genotypes or changes associated with germination time points (Supplemental Fig. S6) and do not offer any indication of this unique, nonuniform localization of oleate.

The spatial localization of four different classes of phospholipids (PE, PA, PI, and PC) was compared in germinating seeds at 36 h postimbibition (Fig. 7A). These analyses also evaluated the distribution of different phospholipid molecular species arising from different combinations of fatty acyl chains. As may be expected from the fact that new cells are being generated as the embryo is undergoing expansion (in contrast to the endosperm, which is being consumed), all four phospholipids are most highly abundant in the embryo. There are contrasting localization patterns depending on the phospholipid class and on the molecular species of the phospholipids. With all

four phospholipids and in both inbreds, the phospholipids with the most unsaturated acyl chains (36:4) show an even distribution between the radicle and the other embryonic tissues. For PI, PC, and PA, the more saturated phospholipids (i.e. in the order 36:3, 36:2, and 36:1) are less concentrated in the radicle compared with other embryonic tissues. In contrast, PE shows a unique localization pattern, with molecules of an even number of double bonds (36:4 and 36:2) being distributed homogenously over the entire embryo, including the radicle, but molecules with an odd

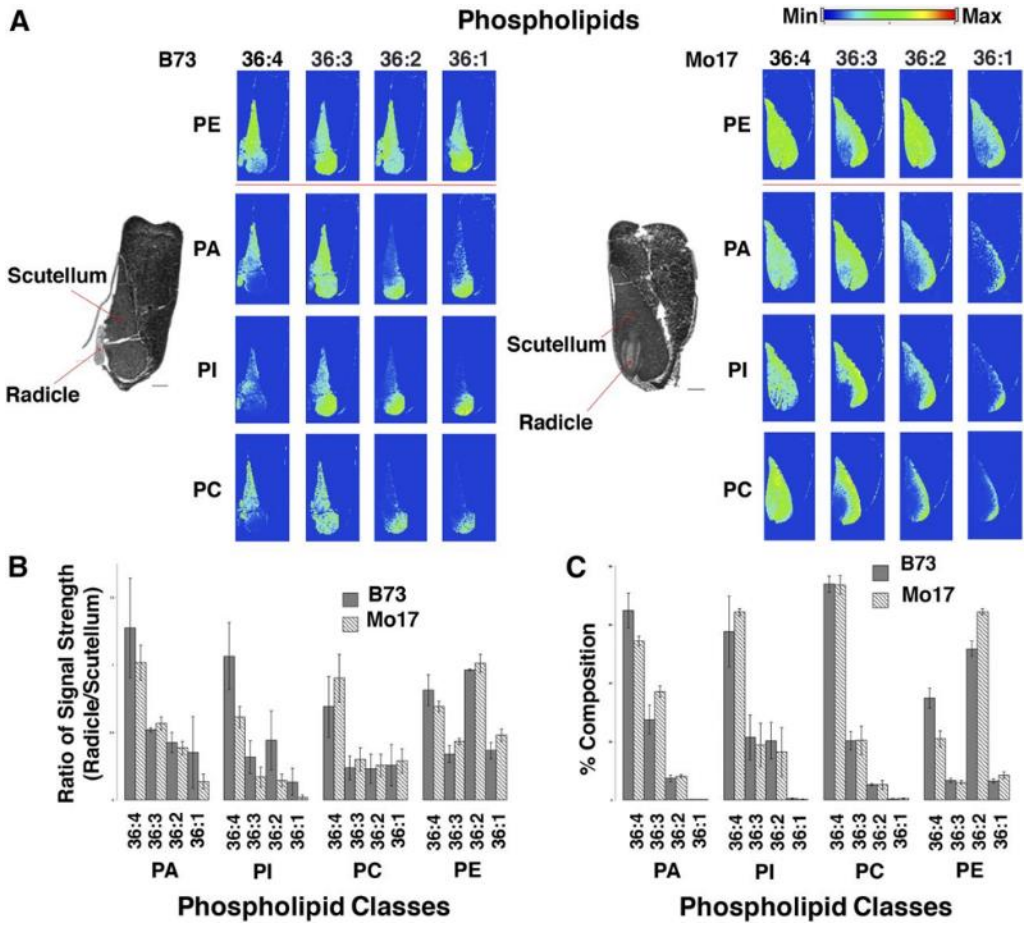


Figure 7 Distribution and species analysis of various phospholipids. **(A)** Distribution of 16 different phospholipid molecular species in germinating seeds (36-h post imbibition). PE, PA, and PI were observed as deprotonated species, $[M-H]^-$, with DAN as the matrix in negative ion mode. PC was observed as a potassium adduct, $[M+K]^+$, with DHB matrix in positive ion mode. To allow direct comparisons, the optical images of the seeds were obtained from the adjacent cross-sections to those used to generate the MS images. Arrows in the radicle and scutellum regions of the embryo indicate the areas used to generate averaged mass spectra, from which semi-quantitative data of the phospholipids were gathered (see B and C). **(B)** Phospholipid ion signals in the radicle region normalized relative to the signals obtained from the scutellum. **(C)** Percentage of each molecular species within each phospholipid class in the radicle region.

number of double bonds (36:3 and 36:1) being less abundant in the developing radicle. These differential localization patterns also were observed at all time points during the germination process (Supplemental Fig. S7).

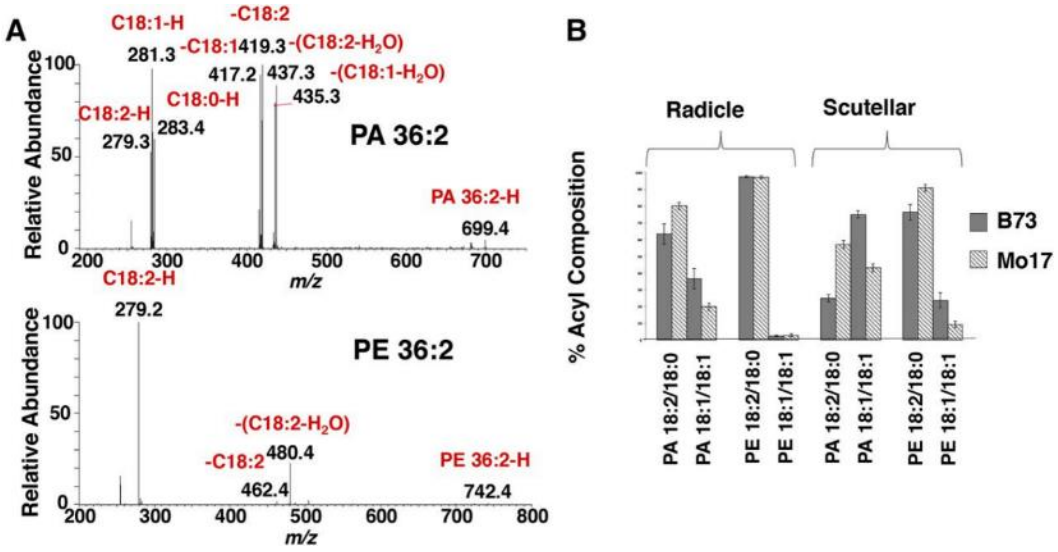


Figure 8. MS/MS analysis of phospholipid species. **(A)** MS/MS spectra of PA 36:2 (top) and PE 36:2 (bottom) obtained from the radicle of an Mo17 inbred seed (24-hr post-imbibition). **(B)** Fatty acyl composition of different molecular species of PA 36:2 and PE 36:2, in the scutellum and radicle region of Mo17 and B73 seeds (24-hr post-imbibition).

To further investigate the nonuniform distribution of phospholipids, quantitative comparison was made from two mass spectra averaged over selected pixels from the radicle and scutellum of seeds developed with a radicle. Figure 7B shows the ratio of each lipid molecular species, comparing signal strength for pixels in the radicle with the signal strength for pixels in the scutellum. This analysis confirms the visual patterns observed from the MS images; namely, the signal strength for the more saturated PA, PI, and PC is lower in the radicle region, whereas the signal strength of PE is lower only for those species with an odd number of double bonds. The unique localization of PE compared with the other phospholipid molecular species may be related to the unusually high abundance of PE 36:2. Figure 7C illustrates these abundance differences among the phospholipids in the radicle region. Whereas the most abundant molecular species are

36:4 for PC, Pam and PI and the more saturated forms of these phospholipids occur at lower levels, 36:2 is the most abundant PE, followed by PE 36:4. With the exception of PE 36:2, which is slightly but statistically higher in Mo17 than in B73 ($P \sim 0.01$), and, accordingly, the lower abundance of PE 36:4 ($P \sim 0.03$), there are no other significant differences between the genotypes.

The differential localization of phospholipids between the radicle and scutellum of the embryo is consistent with the relatively lower abundance of oleate (18:1) in the radicle (Fig. 6). To investigate this possible correlation, we took advantage of the MS/MS capabilities of the MALDI-MSI instrument used in this study and performed MS/MS imaging experiments. In MS/MS, a specific precursor ion can be selected and fragmented, and the mass spectra of fragment ions can be interpreted to extract structural information of the precursor ions. In the current MS/MS imaging experiment, a four-step data acquisition pattern (Supplemental Fig. S8) was used to selectively acquire the MS/MS data for four PA and PE molecular species (PE 36:3, PE 36:2, PA 36:3, and PA 36:2). Figure 8A shows MS/MS spectra obtained from the MS/MS imaging experiment for PA 36:2 and PE 36:2 in the radicle region of a 24-h Mo17 seed (similar results were seen for B73). For these 36:2 lipid species, there are two fatty acyl chain combinations possible: 18:2/18:0 and 18:1/18:1. The MS/MS spectrum of PA 36:2 has each of the three fatty acyl fragments (18:2, 18:1, and 18:0), indicating the presence of both the 18:2/18:0 and 18:1/18:1 combinations. In contrast, MS/MS of PE 36:2 shows only the 18:2 acyl fragment in the radicle of Mo17, indicating that it is composed primarily of the 18:2/18:0 combination. Because MS/MS preferentially induces fragmentation at the sn-2 position for PE and PA species (Hou et al., 2011), the inability to observe the 18:0 acyl fragment could indicate that PE 36:2 has the 18:2 fatty acyl chain located predominantly at the sn-2 position. Parallel MS/MS experiments of

PA 36:3 and PE 36:3 reveal the occurrence of both 18:2 and 18:1 acyl fragments, confirming the occurrence of these two fatty acyl chains in these lipids (Supplemental Fig. S9).

Relative quantification between 18:2/18:0 and 18:1/18:1 was performed from this data set for PA and PE species in the radicle and scutellum, as shown in Figure 8B. Similar to Fgure 7, B and C, in this calculation, MS/MS spectra were averaged from pixels in the radicle and the scutellum and fatty acyl fragments were summed to calculate the proportion of each molecular species. It should be noted, however, that because of the difference in the fragmentation efficiency between the sn-2 and sn-1 positions, this is not an exact quantification of the four lipids. Regardless, these calculations indicate that PE 36:2 (18:2/18:0) is the predominant PE species both in the radicle and scutellum, matching the homogenous images of the distribution of PE 36:2 (Fig. 7A) and of linoleate (Fig. 6) throughout the embryo. In contrast, PA 36:2 (18:1/18:1) occurs at significantly higher levels in the scutellum than in the radicle, consistent with the heterogeneous distribution of oleate among these two tissue types (Fig. 6).

The nonsymmetric distribution of 18:2-containing phospholipids between scutellar and radicle tissues suggests that the supply of 18:2 fatty acid may be differential between these two tissues. We directly tested this hypothesis by assaying the expression of the fatty acid desaturase that converts 18:1 to 18:2. Specifically, we evaluated *FAD2* mRNA levels by qRT-PCR using RNA isolated from the dissected scutellar and radicle tissues. Bioinformatic analysis of maize genome data (Sen et al., 2010; Petryszak et al., 2016) identified six putative *FAD2* genes based upon sequence homology to the Arabidopsis *FAD2* gene (At3g12120); these were annotated as GRMZM2G169240, GRMZM2G169261, GRMZM2G174766, GRMZM2G056252, GRMZM2G161792, and GRMZM2G064701. Initial assays (Supplemental table S5) identified that only GRMZM2G056252 and GRMZM2G064701 are expressed at measurable levels in

germinating embryonic tissue during the initial 36 h after imbibition; the expression of the other four FAD2-like genes was detectable by RT-PCR, but their levels were too low to obtain robust quantitative expression data. However, later in the germination process (at 48 h postimbibition), two of these *FAD2*-like genes were induced (i.e. GRMZM2G169240 and GRMZM2G169261; Supplemental Table S5), but this induction was not relevant to testing of the MSI-generated hypothesis.

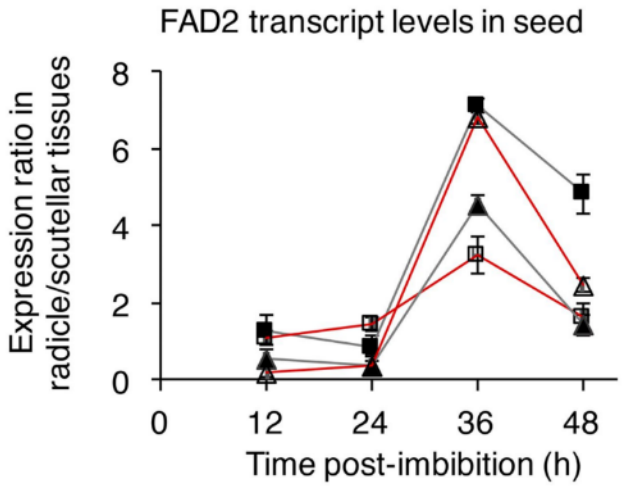


Figure 9. Relative FAD2 expression in radicle and scutellar tissues of germinating maize seeds. FAD2 transcript levels encoded by GRMZM2G056252 (squares) and GRMZM2G064701 (triangles) were determined in RNA isolated from the radicle and scutellar tissues of germinating seeds from the inbreds B73 (gray lines) and Mo17 (red lines) 12 to 48-h post-imbibition. The expression of FAD2 transcripts was calculated by $2^{-(\Delta\Delta Ct)}$, using the ubiquitin mRNA (GenBank accession no. BT018032) as the internal control. The expression of each FAD2 transcript was normalized relative to the level found in the scutellar tissue.

Figure 9 shows the expression profile of the GRMZM2G056252 and GRMZM2G064701 mRNAs over the seed germination time period evaluated by MSI and metabolomics analyses. These data indicate that these two *FAD2* genes are initially expressed at similar levels between the radicle and scutellum, but expression increases subsequently in the radicle by between 3- and 7-fold. This increased expression in the radicle occurs in both inbreds (B73 and Mo17), reaching peak expression at 36 h postimbibition. These data are consistent with the elevated levels of 18:2-containing PE in the radicle.

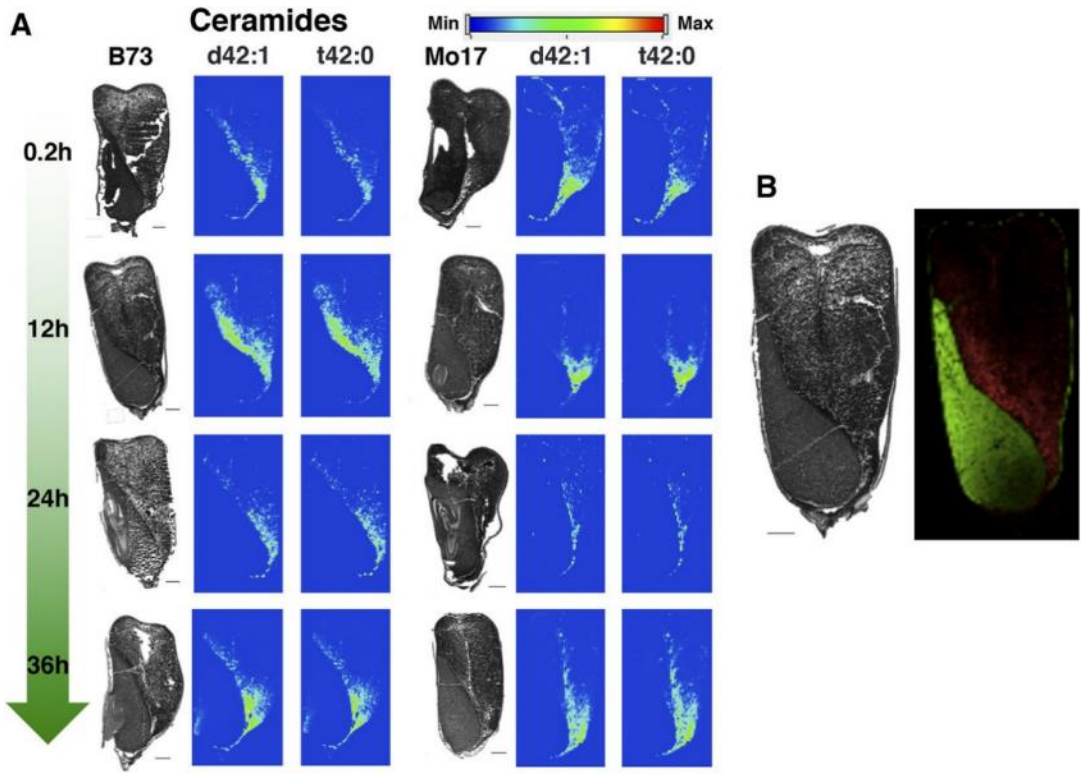


Figure 10. Distribution of ceramide species. **(A)** Distributions of two ceramide molecular species (Cer d42:1, Cer t42:0). Both are detected as protonated species, $[M+H]^+$, in positive ion mode with DHB as the matrix. Bar = 1 mm. **(B)** Combined images of PC 34:2 (green) and ceramide d42:1 (red) in a B73 inbred seed at 12-h post-imbibition. Bar = 1 mm.

Several ceramide (Cer) molecular species also were observed in both inbreds, and their accumulation is persistent throughout the course of germination. Figure 10A shows the distribution of Cer d42:1 and Cer t42:0, and the distributions of other Cer species (Cer d42:2, Cer d40:1, and Cer t40:0) are shown in Supplemental Figure S10. We previously performed MS/MS of Cer d42:1 in a germinated B73 maize seed, confirming its structure (Feenstra et al., 2015), and we also are confident of the assignment of other ceramides based on the accurate mass determinations of the ions, all within 5 ppm mass error; however, MS/MS was not successful for other ceramides due to their low abundances. All ceramides have unique localization patterns that are different from the other lipid classes that were imaged. In both inbreds, and throughout the course of germination, all the detected ceramides are located on the endosperm side of the endosperm-

scutellum boundary. Figure 10B shows this contrasting localization, comparing the distribution of Cer d42:1 with the embryo-specific PC 34:2 relative to the endosperm-scutellum boundary.

Discussion

The process of seed germination is a particularly interesting biological process driven by cellular and metabolic coordination among several spatially distinct compartments to successfully establish the seedling. Because of the agricultural importance of cereal grains, previous metabolomics-based work has focused on these processes during the germination of rice (*Oryza sativa*; Shu et al., 2008) and the malting process for barley (*Hordeum vulgare*; Frank et al., 2011). However, traditional metabolomics analyses often prove laborious and challenging, with limited spatial information. Conversely, MSI presents a straightforward analytical capability that provides high-resolution spatial distribution data for small metabolites (less than 1.5 kD) but at the cost of limited quantitative information and total number of metabolites detected. In this work, a combination of GC- and LC-based metabolomics analyses was combined with a MALDI-MSI approach to study the quantitative metabolite profile data and spatial distribution of metabolites during the germination of seeds of two maize inbreds, Mo17 and B73. These complementary strategies provided a path to investigate both genotypic and developmental differences that occur during germination as well as provide quantitative and finely detailed spatial localization of metabolites.

The genetic diversity within maize is vast, as evidenced recently by rampant structural variation in genic content and copy number variation across maize inbreds (Springer et al., 2009; Lai et al., 2010; Swanson-Wagner et al., 2010; Hirsch et al., 2014). Inbreds B73 and Mo17 were selected in state-sponsored public breeding programs led by the Land Grant institutions Iowa State University and University of Missouri, respectively, in the early 1900s and continuing to today.

B73 is derived from the Stiff Stalk Synthetic population generated and maintained at Iowa State University, and Mo17 was selected from Lancaster Sure Crop material first developed in Lancaster, Pennsylvania (Troyer, 2004). The B73 and Mo17 inbreds share a rich history in the public sector as key founders of U.S. germplasm (Lu and Bernardo, 2001) and the development of agronomically important hybrids, both in the public and private sectors (Troyer, 2009). These important inbreds differ significantly in their genomic structures, both in terms of presence-absence variation in gene content and in the prevalence of single-nucleotide polymorphisms, occurring on average every ~80 bp (Fu et al., 2006; Vroh Bi et al., 2006) based on comparison of published genome sequences(Schnable et al., 2009). The genetic diversity between B73 and Mo17 translates into metabolic, physiological, and phenotypic differences, including differences in germination rates and germination efficiency for inbreds B73 and Mo17 that have been observed under different environmental conditions(Munamava et al., 2004). This genetic diversity between the two inbreds also has led to differences in the metabolome, as was observed in our metabolic assessment of the early germination of maize seedlings. The most significant metabolic difference between germinating B73 and Mo17 seeds was in the enhanced accumulation of free amino acids in Mo17.

The fact that four of the amino acids (i.e. Arg, Asn, Gln, and Lys) showing higher abundance in Mo17 versus B73 are the carriers of reduced nitrogen may indicate that these metabolic processes are of significance in the interaction between carbon and nitrogen metabolism, as the germinating seeds adjust the physiological needs of the emerging seedling from different starting points in the amino acid profiles available in each inbred. Consistent with this hypothesis is the fact that these amino acid profiles are differentially affected by the germination processes, with Thr, Lys, Arg, and Asn occurring at higher levels in Mo17 than in B73 at the earlier stages

of germination (12 and 24 h postimbibition) and Gln levels also increasing in abundance in Mo17 at early stages and in B73 by 36 h postimbibition. Germination efficiency has been shown to be correlated to the differential expression of enzymes involved in Asp-derived metabolism and, in turn, the differential metabolism of specific amino acids (Anzala et al., 2006) as well as nitrogen-derived amino acid metabolism, specifically related to an amino acid-anabolism enzyme, Gln synthetase (Limami et al., 2002). Moreover, quantitative genetic approaches utilizing B73 x Mo17-derived germplasm to dissect the genetic basis for germination phenotypes have identified candidate genes involved in amino acid metabolism (Kollipara et al., 2002).

A second clear metabolome difference between the two inbreds is in the preferential accumulation of small organic acids in B73 (i.e. malate, pyruvate, and fumarate) compared with Mo17. Interestingly, quantitative genetic analyses of recombinant inbred populations derived from B73 and Mo17 parental inbreds have identified quantitative trait loci associated with germination under both optimal and low-temperature conditions. Although numerous candidate genes were identified (more than 3,000), one encodes a malate dehydrogenase (Hu et al., 2016). Collectively, the observed differences in metabolite accumulation between Mo17 and B73 across early germination provide testable hypotheses regarding their impacts on differential germination efficiency and other germination-related traits that have been shown to differ across maize inbreds.

Chromatography-based metabolite profiling generated quantitative data of small molecules, such as sugar monomers, organic acids, amines and amino acids, fatty acids, and sterols, which establish that germinating seeds of the two maize inbreds are readily distinguishable at the level of the expressed metabolome (Fig. 2). While these analyses provided global profiles of the seed metabolomes from which one can generate statistical quantitative correlations among individual metabolite abundances as affected by development and genotype, detailed information

on the localization of the metabolites at the cellular level was lost. Further such analyses of seeds microdissected to separate different seed organs provided compartmentalization information of the metabolites and demonstrated that the metabolome of these different organs vary with respect to the developmental program of germination and is further differentially affected by different genotypes (Fig. 2, C and D). In general, however, even with the microdissected organs, the localization of the metabolites is limited to large, physically separable structures, each of which is made up of a combination of different tissue types with differing metabolic capabilities. Cell type-specific analysis can be done by sorting cells or using laser-capture microdissection, but such analysis is not typically applicable for metabolomic profiling because of the potential metabolite turnover during the laborious sample preparation process. Furthermore, the unknown degree to which these physically separated organs may be cross-contaminated complicates the interpretation of these data. MSI, therefore, offers a convenient means of overcoming the limitations of the chromatography-based metabolite profiling techniques in generating spatially resolved data, with fine-scale localization information in maize seeds at the cellular level.

Combining MALDI-MSI data with chromatography-based analyses allows for more comprehensive coverage of the metabolome beyond that obtainable by either one individually. For example, GC-MS revealed the relative abundance of small sugar molecules (tetroses, pentoses, hexoses, and disaccharide), and analysis of microdissected seed organs indicated that these small sugars are localized primarily to the embryo compared with the endosperm, and they also occur in the pericarp and/or aleurone tissues. MSI was only able to visualize the disaccharide species, and those data are in agreement with the distribution determined from the chromatographic analysis of microdissected organs. However, MSI provided additional information by successfully visualizing the location of large polysaccharides (Hex_5–Hex_9), all of which are localized to the starchy

endosperm of the seed (Fig. 3A). These results are consistent with the expectation that starch breakdown during germination occurs initially by the action of α- and β-amylase secreted from the scutellum and aleurone. These oligosaccharides are broken down to their constituent Glc units and are transported to and taken up by the scutellum, where they are reassembled into Suc, the disaccharide likely visualized by MSI (Bewley, 2001; Nonogaki, 2008). The localization of the presumable Suc disaccharide in the emerging radicle is consistent with the previously reported mobilization of Suc within the scutellum and the subsequent transport to the embryonic axis, specifically to fuel cell elongation in the radicle (Sánchez-Linares et al., 2012).

A similar benefit of combining metabolomic and MSI data is observed when one considers the patterns observed for the different fatty acids and lipid molecular species. GC-MS analysis was limited to being able to detect fatty acids primarily of 16- and 18-carbon chain lengths after derivatization to volatile methyl esters. In contrast, MALDI-MSI not only cannot detect these fatty acids without derivatization but also cannot detect larger nonvolatile lipid molecules such as phospholipids, ceramides, and TAGs. Both MSI and GC-MS showed similar fatty acid profiles, with minor differences in their relative abundances between the two inbreds. However, MSI revealed differential localization of the fatty acids dependent on the degrees of unsaturation. Homogenous distributions were observed within the embryo for palmitate and linoleate, but absences were seen in radicles for oleate and partially for stearate (Fig. 6).

Similar differential patterns of spatial distribution were observed for phospholipids and TAGs depending on their fatty acyl species, in that oleate contributes to the absence in radicles and linoleate contributes to the homogenous embryonic distributions. For example, fully unsaturated TAG 54:6 (18:2/18:2/18:2) is highly homogenous throughout the embryo but becomes absent in the radicle as the degree of saturation increases (Fig. 3B). Phospholipids of PI, PA, and

PC show the same trend as TAGs, with their absence in the radicle being correlated with the degree of saturation (Fig. 7). PE shows apparently different localization from other phospholipids by having PE molecular species with even numbers of unsaturation (36:4 and 36:2) evenly distributed throughout the radicle and embryo. However, MS/MS experiments revealed that this is due to the high radicle content of 18:1/18:1 in other 36:2 phospholipids, while PE 36:2 is mostly composed of 18:2/18:0 (Fig. 8B). These results suggest that the heterogeneous distribution of fatty acids dictates the fatty acyl chain-dependent differential localization of lipids. Such nonuniform distribution has been observed previously in cotton (*Gossypium hirsutum*; Horn et al., 2012) and *Camelina* spp. seeds (Horn et al., 2013); thus, this is not unique to maize but, apparently, is a more general phenomenon. However, the specifics of the distribution are unique for each species, likely owing to the different anatomy and biological properties of these seeds.

Because plant systems use PC as the substrate for fatty acid desaturation, specifically the FAD2 enzyme that generates 18:2 (Bates et al., 2013; Li-Beisson et al., 2013), we evaluated whether this nonuniform distribution of the unsaturated phospholipids may reflect the distribution of this biosynthetic enzyme. We found that only two of the six FAD2-encoding genes of maize are expressed at this stage of seed germination (namely, GRMZM2G056252 and GRMZM2G064701). Moreover, consistent with the hypothesis generated from the MSI data, we found that these two *FAD2* genes are preferentially expressed in the radicle tissue relative to the scuttelar tissue, providing a potential explanation for the enrichment of the 18:2-containing lipids in the former.

Ceramide molecular species detected by MSI showed incredibly unique localizations that were significantly different from any other lipid species detected in this work. By directly comparing with PC molecular species, we found that the ceramides are localized to the endosperm

side of the endosperm/scutellum boundary. Previous work has reported the presence of ceramides in maize (Dietrich et al., 2005) as well as in the bran and endosperm of rice grains (Fujino et al., 1985), supporting our observed localization of ceramides in the endosperm of the maize seeds. Although detailed biochemical understanding of the biological functions that ceramides play in plants is relatively poor in comparison with mammalian and fungal organisms (Kolesnick, 2002; Levy and Futerman, 2010), they are thought to be important for plant growth and defense. Inhibition of their biosynthesis by a mycotoxin (Williams et al., 2007) or mutations in the biosynthesis pathway lead to programmed cell death and disease states in plants (Lynch and Dunn, 2004; Markham et al., 2013), whereas induction of their biosynthesis via the overexpression of ceramide synthases leads to impacts in biomass accumulation, growth, and mycotoxin resistance (Luttgeharm et al., 2015).

In conclusion, this study successfully demonstrated the combined use of chromatography-based metabolite profiling and MALDI-MSI-based chemical imaging to elucidate the metabolite profile and spatial distribution of metabolites in the complex processes of seed germination in maize. The traditional metabolomics approaches offered comprehensive metabolite coverage of small molecules, which demonstrated that there are significant differences in the metabolite profiles between inbreds during germination as well as between different seed organ/tissue types of the same inbred. Conversely, MALDI-MSI offers more detailed spatial information on the distribution of the metabolites, especially in terms of unusual localizations of lipids. This study demonstrates the synergy that can be gained by combining these strategies to reveal new insights into complex biological processes that integrate the capabilities of different cellular compartments.

Materials and Methods

Seed germination and harvest

Seeds of maize (*Zea mays*) inbreds B73 and Mo17 were imbibed and germinated as described previously (Liu et al., 2013). Briefly, seeds were placed in deionized water and imbibed while shaking at 300 rpm for 10 min. These seeds were then placed embryo side down on moist filter paper in petri dishes. Plates were incubated in a greenhouse under a diurnal cycle of 15 h of light and 9 h of dark at 27°C and 24°C, respectively.

Germinating seeds were collected at four time points, 0.2, 12, 24, and 36 h, with the 0.2-h time point corresponding to those taken immediately after 10 min of imbibition. At each time point, seeds were collected that appeared similar in terms of their developmental stage. Three seeds chosen for MSI were sectioned longitudinally, immediately flash frozen in liquid nitrogen, and stored at −80°C. From another batch of seeds grown in identical conditions, nine seeds were chosen at each time point for total metabolite and amino acid analyses. Whole seeds were frozen in liquid nitrogen and dried using a vacuum lyophilizer (Labconco); then, three seeds were combined as one biological replicate and pulverized using a Mixer Mill 301(Retsch) in 2-mL Eppendorf tubes for total metabolite analysis. Companion sets of germinating seeds of similar shapes were collected at the 12- and 36-h time points, microdissected to separate the pericarp and/or aleurone (these two layers could not be separated), tip cap, embryo, and endosperm, and then the same lyophilization/pulverization process described above was followed for each dissected tissue.

Cryosectioning

Each seed was removed from −80°C storage and quickly placed cut side down into a plastic cryo-mold (Electron Microscopy Sciences). A warm, 10% (w/v) gelatin solution was poured into the cryo-mold such that the solution completely covered the seed. The mold was then held over

liquid nitrogen until the gelatin became mostly opaque (approximately one-half of the gelatin) and then transferred to a −20°C freezer to complete the freezing process.

The frozen gelatin block was removed from the cryo-mold and mounted on a cryostat chuck. The seed was then cut down to the tissue region of interest (e.g. the embryo axis containing the radicle, if the radicle was present in the seed) using a Leica CM1520 cryostat (Leica Biosystems) until the embryonic tissue was visible, following which consecutive 10-µm tissue sections were collected on adhesive tape windows (Leica Biosystems) to preserve the structural integrity of the tissue sections. The adhesive windows were then taped face up onto chilled glass slides and stored at −80°C. A total of 25 sections were taken from each seed, and a subset was subjected to either microscopic imaging or MSI analysis.

Microscopy and MALDI imaging preparation

Microscope images of four sections (nos. 5, 11, 17, and 22 of the 25 collected) were obtained using an Olympus SAH-10 stereomicroscope (Olympus Scientific Solutions) with an AxioCam HRC and Axio Vision software (Carl Zeiss) to provide finely detailed images. From these images, the highest quality section was selected based on the clear visualization of developmental stage and section integrity. The sections nearest to this section were chosen for MSI. The sections selected for MSI were placed on a chilled aluminum plate and then allowed to slowly equilibrate to room temperature under vacuum, preventing condensation and water-soluble metabolite delocalization.

Matrix application was performed via sublimation-vapor deposition as described previously (Hankin et al., 2007). Briefly, the sample was attached to the bottom of the sublimation flask condenser, 300 mg of matrix was added to the bottom of the flask, and the flask was sealed and evacuated to ~20 mtorr. The condenser was cooled with a dry ice/acetone slurry. After cooling

the condenser, the flask was placed into a preheated heating mantle and maintained at a constant temperature (140°C for DAN and DHB and 170°C for 9-AA) for 3 to 6 min until matrix deposition was visible on the tissue surface. After matrix application, tape windows were transferred to a MALDI plate and inserted into the mass spectrometer for MSI.

MALDI-MS Imaging

MSI data were acquired using a MALDI-linear ion trap-Orbitrap mass spectrometer (MALDI-LTQ-Orbitrap Discovery; Thermo Finnigan) modified to use an external diode-laser pumped frequency-tripled Nd:YAG laser (355 nm, UVFQ series; Elforlight). For DAN and 9AA in negative ion mode, Orbitrap MS data were acquired using a 100-μm raster size for the *m/z* range of 50 to 1,000. For DHB in positive ion mode, each raster step (100 μm) was broken down into two spiral steps (50 μm each), and Orbitrap MS scans were performed for *m/z* 50 to 800 and *m/z* 600 to 1,600, respectively. For all imaging runs, the laser spot size was ~40 μm estimated from the burn mark on thin DHB film, except the high-spatial-resolution experiment in Figure 5, where the laser spot size was ~10 μm adjusted using a beam expander in the optical beam path. Laser pulse energy and number of shots were optimized for each matrix.

Data analysis and image generation were performed using ImageQuest and Xcalibur software (Thermo). Image normalization was performed for some images and is noted in the figure legends. Mass values and tolerances used to generate images are summarized in Supplemental Table S1. For some metabolites, MS images were generated after normalization to the total ion count at each pixel (defined by the raster step) to minimize spot-to-spot variation, and the same color scale is used regardless of genotype or germination time point. For most of the metabolites, however, MS images are produced in absolute ion scales without normalization, due to the fact that normalization can distort the images for low-abundance ions, and the color scheme was

adjusted arbitrarily for each image to obtain an image with the highest contrast to display where the metabolite is localized without accentuating background noise.

Separate MS/MS experiments were performed using collision-induced dissociation of the linear ion trap. For *m/z* 193.051, a collision energy of 100 (arbitrary unit) was used with an activation time of 30 ms and an isolation mass window of 0.8 D. For phospholipid MS/MS, a collision energy of 45 was used with an activation time of 30 ms and an isolation mass window of 1 D.

Total Metabolites Analysis

Total metabolite extractions were carried out as described previously (Schmidt et al., 2011). Extracts were prepared from ~20 mg of lyophilized and pulverized whole seeds and dissected seed fractions. Each sample was spiked with two internal standards (25 μg of ribitol and 25 μg of nonadecanoic acid for polar and nonpolar fractions, respectively), and 0.35 mL of hot methanol (60°C) was added and incubated at the same temperature for 10 min, followed by sonication for 10 min at full power. To this slurry, 0.35 mL of chloroform and 0.3 mL of water were added, and the mixture was vortexed for 1 to 3 min. After centrifugation for 5 min at 13,000*g*, 200 μL of the upper phase (polar fraction) and 200 μL of the lower phase (nonpolar fraction) were removed separately into 2-mL GC-MS vials and dried in a Speed-Vac concentrator (model SVC 100H; Savant). Both polar and nonpolar extracts were analyzed for whole-seed samples, and only the polar fractions were analyzed for microdissected samples.

Double derivatization via methoximation and silyation was performed to protect ketone functional groups and to increase the volatility of the compounds, respectively. For methoximation, 50 μL of 20 mg mL^{-1} methoxyamine hydrochloride dissolved in dry pyridine was added, and the reaction was conducted at 30°C for 1.5 h with continuous shaking. Silylation was

performed by the addition of 70 µL of *N*,*O*-bis(trimethylsilyl)trifluoroacetamide with 1% trimethylchlorosilane and incubated at 65°C for 30 min. One microliter of the derivatized samples was injected for GC-MS in splitless mode. GC-MS analysis was performed using an Agilent 6890 gas chromatograph interfaced to an Agilent 5973 quadrupole mass spectrometer with an HP-5ms column (30 m × 0.25 mm × 0.25 µm; Agilent). The temperature gradient was programmed from 70°C to 320°C at 5°C min^{-1} with helium flow rate at 1 mL min^{-1} and inlet temperature at 280°C. EI-MS ionization energy was set to 70 eV, and the interface temperature was 280°C.

Amino acid analysis

Amino acid analysis was carried out following a method similar to that described previously by Guan et al. (2015). About 10 mg of lyophilized maize seed powder was extracted with 1 mL of hot water at 85°C for 30 min and spiked with 10 µM butylamine internal standard. The samples were centrifuged briefly for 5 min, passed through a syringe filter with a 0.2-µm pore size, and subjected to LC-fluorescence amino acid analysis with precolumn derivatization of primary amino acids with OPA.

LC-fluorescence analysis was performed using an Agilent 1110 HPLC device with a fluorescence detector with a Hypersil ODS C18 reverse-phase column (250 mm × 4 mm, 5 µm; Thermo). Excitation and emission wavelengths for fluorescence detection were set at 337 and 454 nm, respectively. The solvent systems used were buffer A (10% methanol in 10 mM NaH_2PO_4, pH 7.3), buffer B (80% methanol in 10 mM NaH_2PO_4, pH 7.3), and OPA solution (12.25 mg of OPA in 312.5 µL of methanol, 6 mL of 0.4 M borate buffer, and 19.22 µL of mercaptoethanol). The precolumn OPA derivatization reaction was achieved using an in-loop-reaction program. The solvent program was a linear gradient from 100% buffer A to 100% buffer B in 46 min at a flow rate of 1 mL min^{-1}. The amino acid standard mixture included butylamine and 18 of the 20

proteogenic amino acids at the concentration of 10 mM each. Cys and Pro cannot be analyzed by this method and were not included as standards. The amino acid standard mixture was used to calculate the response factors for the detector, and butylamine was used to generate a standard curve.

Data Analysis

Three biological replicates were used for GC or LC analysis of the extracts. Those compounds that were observed in two or three data sets were subjected to statistical analysis and reported here. For those compounds observed only in two data sets, the missing values were replaced by one-third of the estimated minimum value. Log_2-ratio plots were calculated for each metabolite as described previously (Quanbeck et al., 2012). The raw data and averages and SE calculations were determined from three biological replicates and are available in the PMR database and Supplemental Tables S2 to S4, respectively. Student's *t* test was used to calculate *P* values.

The GC-MS data files were deconvoluted and searched against an in-house MS library and the NIST 14 Mass Spectral Library using NIST AMDIS software (Stein, 1999). Total metabolite analysis results obtained with GC-MS and amino acid analysis from LC fluorescence are publicly available in the PMR database (http://metnetdb.org/PMR/; Hur et al., 2013). They are available from the site in the name Seed Germination under the species *Zea mays*.

RNA Extraction and Quantitative PCR

Embryos were microdissected from germinating seeds at different times after imbibition, and the scutellar tissue was separated from the radicle. The flash-frozen tissues were pulverized, and RNA was extracted (Wang et al., 2012). The isolated RNA preparations were treated with DNase (Ambion), and 8 μg of RNA was reverse transcribed using Double Primed RNA to cDNA EcoDry Premix (Clontech). *FAD2* gene sequences were identified using the Gramene

(www.gramene.org) and Maize Genome (www.maizegdb.org) databases. Primer pairs for qRT-PCR were designed and computationally tested for specificity using Primer-BLAST (Ye et al., 2012). PCR was performed with the StepOnePlus Real-Time PCR System (Applied Biosystems). The PCR mixture contained cDNA template, primer mix, and SYBR Select Master Mix (Applied Biosystems) in a final volume of 20 μL. The annealing temperature in the PCR was 58°C using gene-specific oligonucleotide primers (Supplemental Table S6). The Ct ($2^{-\Delta\Delta Ct}$) method (Livak and Schmittgen, 2001) was used to analyze the comparative expression of each FAD2 mRNA, using the expression of the ubiquitin mRNA (GenBank accession no. BT018032) as the internal control.

Accession Numbers

Sequence data from this article can be found in the MaizeGDB/zmGDB data libraries under the following accession numbers: GRMZM2G169240, GRMZM2G169261, GRMZM2G174766, GRMZM2G056252, GRMZM2G161792, and GRMZM2G064701.

Acknowledgments

Metabolomics analyses were conducted at the Iowa State University W. M. Keck Metabolomics Research Laboratory, and we acknowledge Ann M. Perera for her expert advice.

Authors Contribution

Y.J.L. conceived the idea; A.D.F., M.D.Y-N., B.J.N. and Y.J.L. designed the experiments; B.J.N. and Y.J.L. supervised experiments; A.D.F. and L.E.A. performed most of the experiments and data analysis; Z.S. and A.R.K. performed some experiments and data analysis; A.D.F., L.E.A., M.D.Y-N., B.J.N. and Y.J.L. wrote the manuscript.

References

Amstalden van Hove ER, Smith DF, Heeren RM (2010) A concise review of mass spectrometry imaging. J Chromatogr A **1217:** 3946–3954

Anzala F, Morère-Le Paven M-C, Fournier S, Rondeau D, Limami AM (2006) Physiological and molecular aspects of aspartate-derived amino acid metabolism during germination and post-germination growth in two maize genotypes differing in germination efficiency. Journal of experimental botany **57:** 645-653

Bates PD, Stymne S, Ohlrogge J (2013) Biochemical pathways in seed oil synthesis. Current opinion in plant biology **16:** 358-364

Bernier AM, Ballance GM (1993) Induction and secretion of alpha-amylase, (1- 3),(1- 4)-beta-glucanase and (1- 3)-beta-glucanse activities in gibberellic-acid and CaCl2-treated half seeds and aleurones of wheat. Cereal Chemistry **70:** 127-132

Bewley JD (1997) Seed germination and dormancy. The plant cell **9:** 1055

Bewley JD (2001) Seed Germination and Reserve Mobilization. eLS

Bewley JD, Bradford K, Hilhorst H (2012) Seeds: physiology of development, germination and dormancy. Springer Science & Business Media

Chrispeels MJ, Varner J (1967) Hormonal control of enzyme synthesis: on the mode of action of gibberellic acid and abscisin in aleurone layers of barley. Plant physiology **42:** 1008-1016

Dante RA, Larkins BA, Sabelli PA (2015) Cell cycle control and seed development. Advances in Seed Biology**:** 21

Dietrich CR, Perera M, D Yandeau-Nelson M, Meeley RB, Nikolau BJ, Schnable PS (2005) Characterization of two GL8 paralogs reveals that the 3-ketoacyl reductase component of fatty acid elongase is essential for maize (Zea mays L.) development. The Plant Journal **42:** 844-861

Feenstra AD, Hansen RL, Lee YJ (2015) Multi-matrix, dual polarity, tandem mass spectrometry imaging strategy applied to a germinated maize seed: toward mass spectrometry imaging of an untargeted metabolome. Analyst **140:** 7293-7304

Fincher GB (1989) Molecular and cellular biology associated with endosperm mobilization in germinating cereal grains. Annual review of plant biology **40:** 305-346

Firenzuoli A, Vanni P, Ramponi G, Baccari V (1968) Changes in enzyme levels during germination of seeds of Triticum durum. Plant physiology **43:** 260-264

Frank T, Scholz B, Peter S, Engel K-H (2011) Metabolite profiling of barley: Influence of the malting process. Food chemistry **124:** 948-957

Fu Y, Wen T-J, Ronin YI, Chen HD, Guo L, Mester DI, Yang Y, Lee M, Korol AB, Ashlock DA (2006) Genetic dissection of intermated recombinant inbred lines using a new genetic map of maize. Genetics **174:** 1671-1683

Fujino Y, Ohnishi M, Ito S (1985) Molecular species of ceramide and mono-, di-, tri-, and tetraglycosylceramide in bran and endosperm of rice grains. Agricultural and biological chemistry **49:** 2753-2762

Guan X, Chen H, Abramson A, Man H, Wu J, Yu O, Nikolau BJ (2015) A phosphopantetheinyl transferase that is essential for mitochondrial fatty acid biosynthesis. The Plant Journal **84:** 718-732

Hankin JA, Barkley RM, Murphy RC (2007) Sublimation as a method of matrix application for mass spectrometric imaging. Journal of the American Society for Mass Spectrometry **18:** 1646-1652

He M, Zhu C, Dong K, Zhang T, Cheng Z, Li J, Yan Y (2015) Comparative proteome analysis of embryo and endosperm reveals central differential expression proteins involved in wheat seed germination. BMC plant biology **15:** 1

Hirsch CN, Foerster JM, Johnson JM, Sekhon RS, Muttoni G, Vaillancourt B, Peñagaricano F, Lindquist E, Pedraza MA, Barry K (2014) Insights into the maize pan-genome and pan-transcriptome. The Plant Cell **26:** 121-135

Horn PJ, Korte AR, Neogi PB, Love E, Fuchs J, Strupat K, Borisjuk L, Shulaev V, Lee Y-J, Chapman KD (2012) Spatial mapping of lipids at cellular resolution in embryos of cotton. The Plant Cell **24:** 622-636

Horn PJ, Silva JE, Anderson D, Fuchs J, Borisjuk L, Nazarenus TJ, Shulaev V, Cahoon EB, Chapman KD (2013) Imaging heterogeneity of membrane and storage lipids in transgenic Camelina sativa seeds with altered fatty acid profiles. The Plant Journal **76:** 138-150

Hou W, Zhou H, Khalil MB, Seebun D, Bennett SA, Figeys D (2011) Lyso-form fragment ions facilitate the determination of stereospecificity of diacyl glycerophospholipids. Rapid Communications in Mass Spectrometry **25:** 205-217

Hu S, Lübberstedt T, Zhao G, Lee M (2016) QTL Mapping of Low-Temperature Germination Ability in the Maize IBM Syn4 RIL Population. PloS one **11:** e0152795

Hur M, Campbell AA, Almeida-de-Macedo M, Li L, Ransom N, Jose A, Crispin M, Nikolauh BJ, Wurtele ES (2013) A global approach to analysis and interpretation of metabolic data for plant natural product discovery. Natural Product Reports **30:** 565-583

Ingle J, Beevers L, Hageman R (1964) Metabolic changes associated with the germination of corn. I. Changes in weight and metabolites and their redistribution in the embryo axis, scutellum, and endosperm. Plant physiology **39:** 735

Kaspar S, Peukert M, Svatos A, Matros A, Mock HP (2011) MALDI-imaging mass spectrometry–An emerging technique in plant biology. Proteomics **11:** 1840-1850

Kolesnick R (2002) The therapeutic potential of modulating the ceramide/sphingomyelin pathway. The Journal of clinical investigation **110:** 3-8

Kollipara KP, Saab IN, Wych RD, Lauer MJ, Singletary GW (2002) Expression profiling of reciprocal maize hybrids divergent for cold germination and desiccation tolerance. Plant Physiology **129:** 974-992

Korte AR, Yandeau-Nelson MD, Nikolau BJ, Lee YJ (2015) Subcellular-level resolution MALDI-MS imaging of maize leaf metabolites by MALDI-linear ion trap-Orbitrap mass spectrometer. Analytical and bioanalytical chemistry **407:** 2301-2309

Lai J, Li R, Xu X, Jin W, Xu M, Zhao H, Xiang Z, Song W, Ying K, Zhang M (2010) Genome-wide patterns of genetic variation among elite maize inbred lines. Nature genetics **42:** 1027-1030

Larson G, Piperno DR, Allaby RG, Purugganan MD, Andersson L, Arroyo-Kalin M, Barton L, Vigueira CC, Denham T, Dobney K (2014) Current perspectives and the future of domestication studies. Proceedings of the National Academy of Sciences **111:** 6139-6146

Lee YJ, Perdian DC, Song Z, Yeung ES, Nikolau BJ (2012) Use of mass spectrometry for imaging metabolites in plants. The Plant Journal **70:** 81-95

Leonova S, Grimberg Å, Marttila S, Stymne S, Carlsson AS (2010) Mobilization of lipid reserves during germination of oat (Avena sativa L.), a cereal rich in endosperm oil. Journal of experimental botany **61:** 3089-3099

Levy M, Futerman AH (2010) Mammalian ceramide synthases. IUBMB life **62:** 347-356

Li-Beisson Y, Shorrosh B, Beisson F, Andersson MX, Arondel V, Bates PD, Baud S, Bird D, Debono A, Durrett TP, et al (2013) Acyl-lipid metabolism. The Arabidopsis Book **11:** e0161, doi/10.1199/tab.0161

Limami AM, Rouillon C, Glevarec G, Gallais A, Hirel B (2002) Genetic and physiological analysis of germination efficiency in maize in relation to nitrogen metabolism reveals the importance of cytosolic glutamine synthetase. Plant Physiology **130:** 1860-1870

Liu W-Y, Chang Y-M, Chen SC-C, Lu C-H, Wu Y-H, Lu M-YJ, Chen D-R, Shih AC-C, Sheue C-R, Huang H-C (2013) Anatomical and transcriptional dynamics of maize embryonic leaves during seed germination. Proceedings of the National Academy of Sciences **110:** 3979-3984

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-D D C(T)) method. Methods **25:** 402–408

Lu H, Bernardo R (2001) Molecular marker diversity among current and historical maize inbreds. Theoretical and Applied Genetics **103:** 613-617

Luttgeharm KD, Chen M, Mehra A, Cahoon RE, Markham JE, Cahoon EB (2015) Overexpression of Arabidopsis ceramide synthases differentially affects growth, sphingolipid metabolism, programmed cell death, and mycotoxin resistance. Plant physiology **169:** 1108-1117

Lynch DV, Dunn TM (2004) An introduction to plant sphingolipids and a review of recent advances in understanding their metabolism and function. New phytologist **161:** 677-702

Markham JE, Lynch DV, Napier JA, Dunn TM, Cahoon EB (2013) Plant sphingolipids: function follows form. Current opinion in plant biology **16:** 350-357

McDonnell LA, Heeren R (2007) Imaging mass spectrometry. Mass spectrometry reviews **26:** 606-643

Munamava MR, Goggi AS, Pollak L (2004) Seed quality of maize inbred lines with different composition and genetic backgrounds. Crop science **44:** 542-548

Nonogaki H (2008) Seed germination and reserve mobilization. eLS DOI/ 10.1002/9780470015902.a0002047.pub2

Petryszak R, Keays M, Tang YA, Fonseca NA, Barrera E, Burdett T, Füllgrabe A, Fuentes AMP, Jupp S, Koskinen S, et al (2016) Expression Atlas update: an integrated database of gene and protein expression in humans, animals and plants. Nucleic Acids Res **44:** D746–D752

Quanbeck SMM, Brachova L, Campbell AA, Guan X, Perera A, He K, Rhee SY, Bais P, Dickerson J, Dixon P (2012) Metabolomics as a hypothesis-generating functional genomics tool for the annotation of Arabidopsis thaliana genes of "unknown function". Frontiers in plant science **3:** 15

Römisch-Margl L, Spielbauer G, Schützenmeister A, Schwab W, Piepho H-P, Genschel U, Gierl A (2010) Heterotic patterns of sugar and amino acid components in developing maize kernels. Theoretical and Applied Genetics **120:** 369-381

Sánchez-Linares L, Gavilanes-Ruíz M, Díaz-Pontones D, Guzmán-Chávez F, Calzada-Alejo V, Zurita-Villegas V, Luna-Loaiza V, Moreno-Sánchez R, Bernal-Lugo I, Sánchez-Nieto S (2012) Early carbon mobilization and radicle protrusion in maize germination. Journal of experimental botany**:** ers130

Schmidt MA, Barbazuk WB, Sandford M, May G, Song Z, Zhou W, Nikolau BJ, Herman EM (2011) Silencing of soybean seed storage proteins results in a rebalanced protein

composition preserving seed protein content without major collateral changes in the metabolome and transcriptome. Plant Physiology **156:** 330-345

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA (2009) The B73 maize genome: complexity, diversity, and dynamics. science **326:** 1112-1115

Sen TZ, Harper LC, Schaeffer ML, Andorf CM, Seigfried TE, Campbell DA, Lawrence CJ (2010) Choosing a genome browser for a Model Organism Database: surveying the maize community. Database 2010: baq007

Shu X-L, Frank T, Shu Q-Y, Engel K-H (2008) Metabolite profiling of germinating rice seeds. Journal of agricultural and food chemistry **56:** 11612-11620

Springer NM, Ying K, Fu Y, Ji T, Yeh C-T, Jia Y, Wu W, Richmond T, Kitzman J, Rosenbaum H (2009) Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. PLoS Genet **5:** e1000734

Stein SE (1999) An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. Journal of the American Society for Mass Spectrometry **10:** 770-781

Swanson-Wagner RA, Eichten SR, Kumari S, Tiffin P, Stein JC, Ware D, Springer NM (2010) Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. Genome research **20:** 1689-1699

Toue S, Sugiura Y, Kubo A, Ohmura M, Karakawa S, Mizukoshi T, Yoneda J, Miyano H, Noguchi Y, Kobayashi T (2014) Microscopic imaging mass spectrometry assisted by on-tissue chemical derivatization for visualizing multiple amino acids in human colon cancer xenografts. Proteomics **14:** 810-819

Troyer AF (2004) Persistent and popular germplasm in seventy centuries of corn evolution. Corn: Origin, History, and Production. John Wiley & Sons, Hoboken, NJ**:** 133-231

Troyer AF (2009) Development of hybrid corn and the seed corn industry. *In* Handbook of maize. Springer, pp 87-114

Vroh Bi I, McMullen M, Sanchez-Villeda H, Schroeder S, Gardiner J, Polacco M, Soderlund C, Wing R, Fang Z, Coe E (2006) Single nucleotide polymorphisms and insertion–deletions for genetic markers and anchoring the maize fingerprint contig physical map. Crop science **46:** 12-21

Wang G, Wang G, Zhang X, Wang F, Song R (2012) Isolation of high quality RNA from cereal seeds containing high levels of starch. Phytochem Anal **23:** 159–163

Williams LD, Glenn AE, Zimeri AM, Bacon CW, Smith MA, Riley RT (2007) Fumonisin disruption of ceramide biosynthesis in maize roots and the effects on plant development

and Fusarium verticillioides-induced seedling disease. Journal of agricultural and food chemistry **55:** 2937-2946

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics **13:** 134

Zavalin A, Yang J, Caprioli R (2013) Laser beam filtration for high spatial resolution MALDI imaging mass spectrometry. Journal of the American Society for Mass Spectrometry **24:** 1153-1156

Zeeman SC, Delatte T, Messerli G, Umhang M, Stettler M, Mettler T, Streb S, Reinhold H, Kötting O (2007) Starch breakdown: recent discoveries suggest distinct pathways and novel mechanisms. *In*, Vol 34. CSIRO Publishing, Functional Plant Biology, pp 465-473

Zhang M, Fan J, Taylor DC, Ohlrogge JB (2009) DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. The Plant Cell **21:** 3885-3901

Supplemental Materials

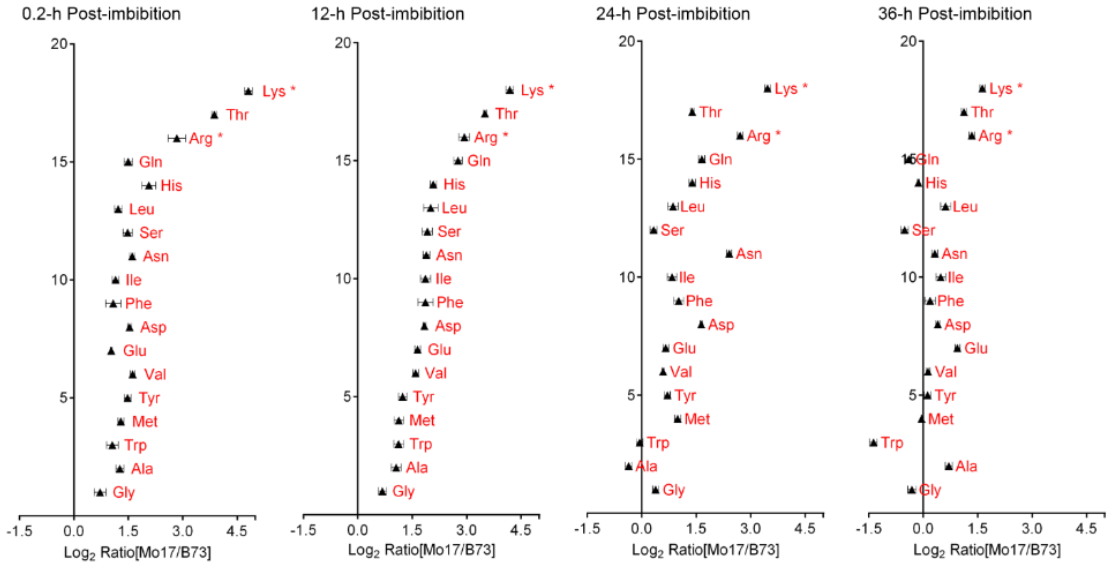


Fig. S1. Log-ratio plot comparison of the amino acid metabolomes of Mo17 and B73 maize inbred seeds. The x-axis plots log-transformed relative abundance ratio of each metabolite in Mo17 vs B73. The y-axis plots the individual metabolites (18 analytes), and the order of the metabolites on the y-axis is identical in all plots and the order is from the lowest to the highest value on the x-axis as determined for the 12-h post-imbibition time point. The asterisks denote those metabolites that were also analyzed by MSI

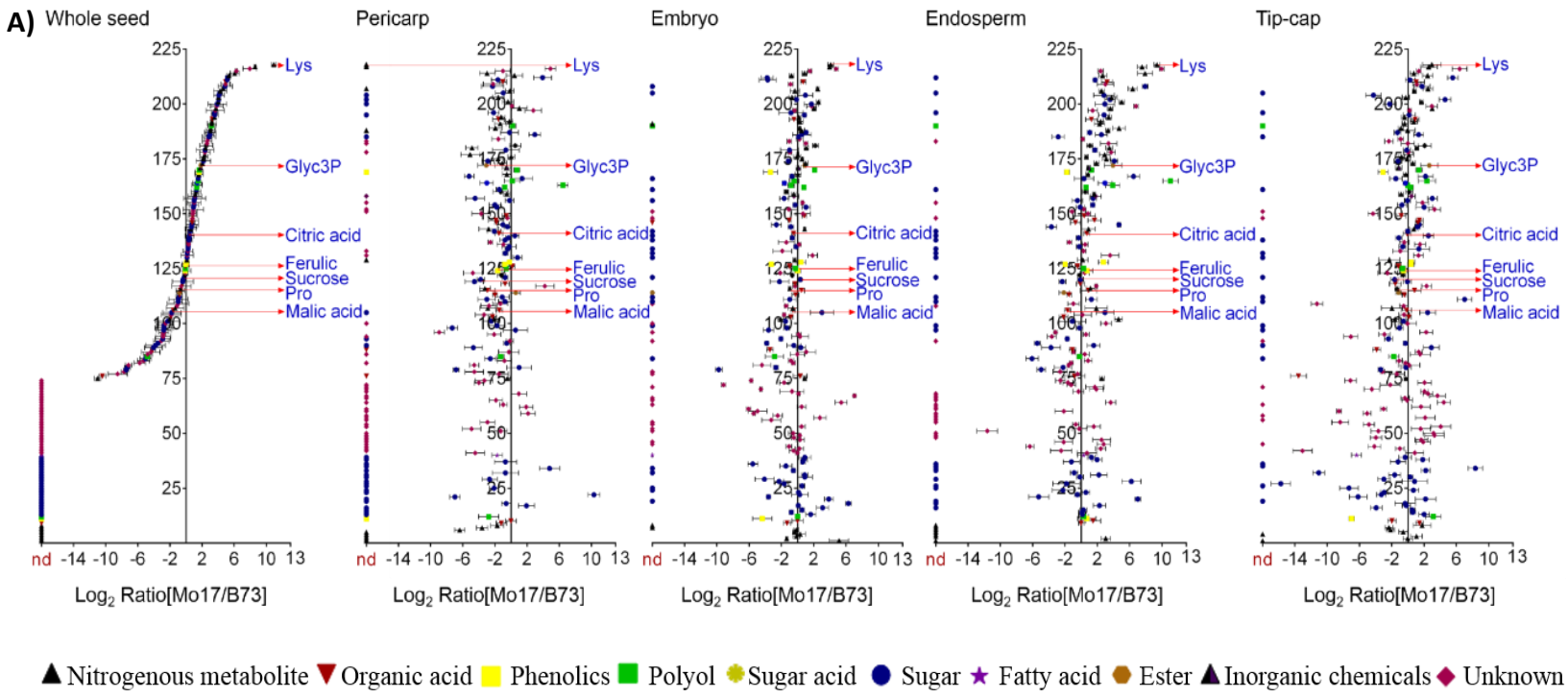


Fig. S2. Log-ratio plot comparison of the metabolome of different micro-dissected organs of Mo17 and B73 maize inbred seeds at 12-h (A) and 36-h (B) post-imbibition. The x-axis plots log-transformed relative abundance ratio of each metabolite in Mo17 vs B73. The y-axis plots the individual metabolites (218 analytes, 75 chemically defined), and the order of the metabolites on the y-axis is identical in all plots and the order is from the lowest to the highest value on the x-axis as determined for the 12-h post-imbibition time point. The arrows denote those metabolites that were also analyzed by MSI. Glyc3P = Glycerol 3-phosphate. "**nd"** on x-axis indicates metabolites whose abundances are indistinguishable between the two inbred lines.

B)

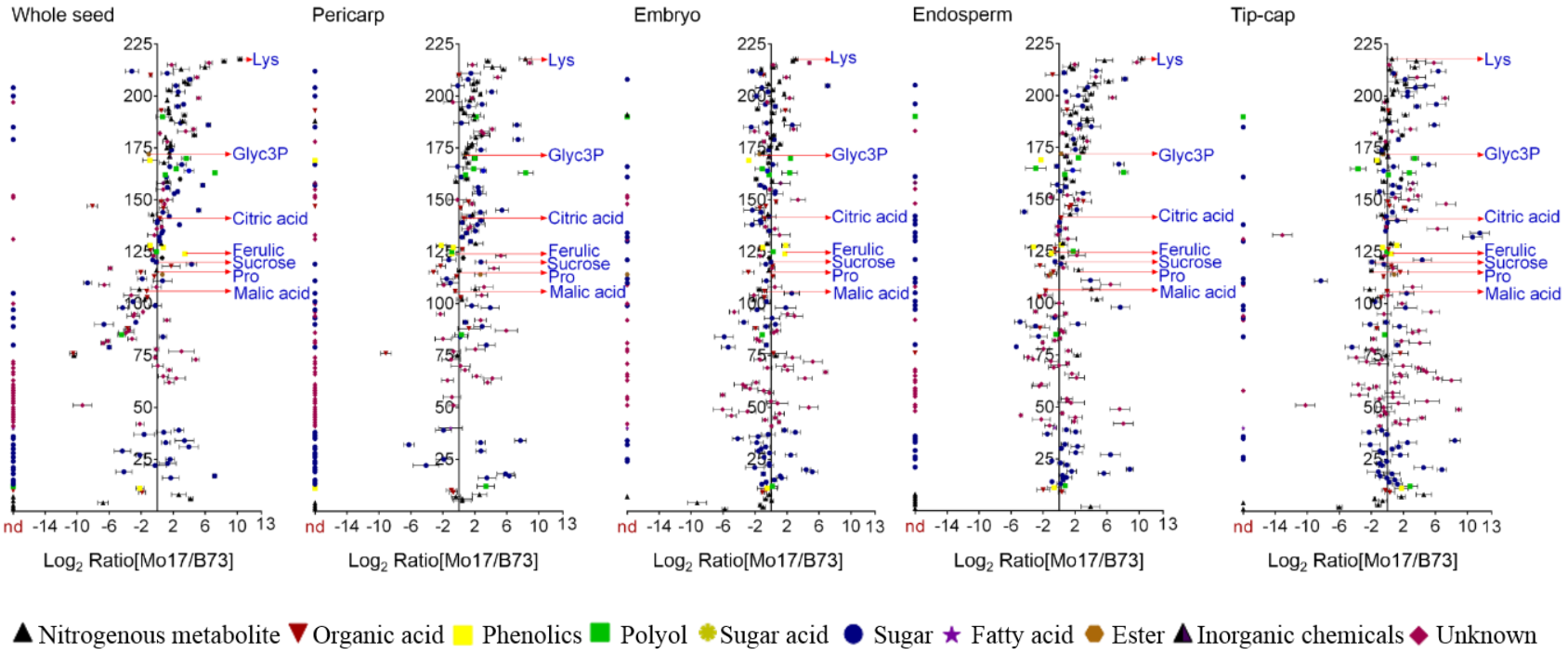


Fig. S2. (cont'd)

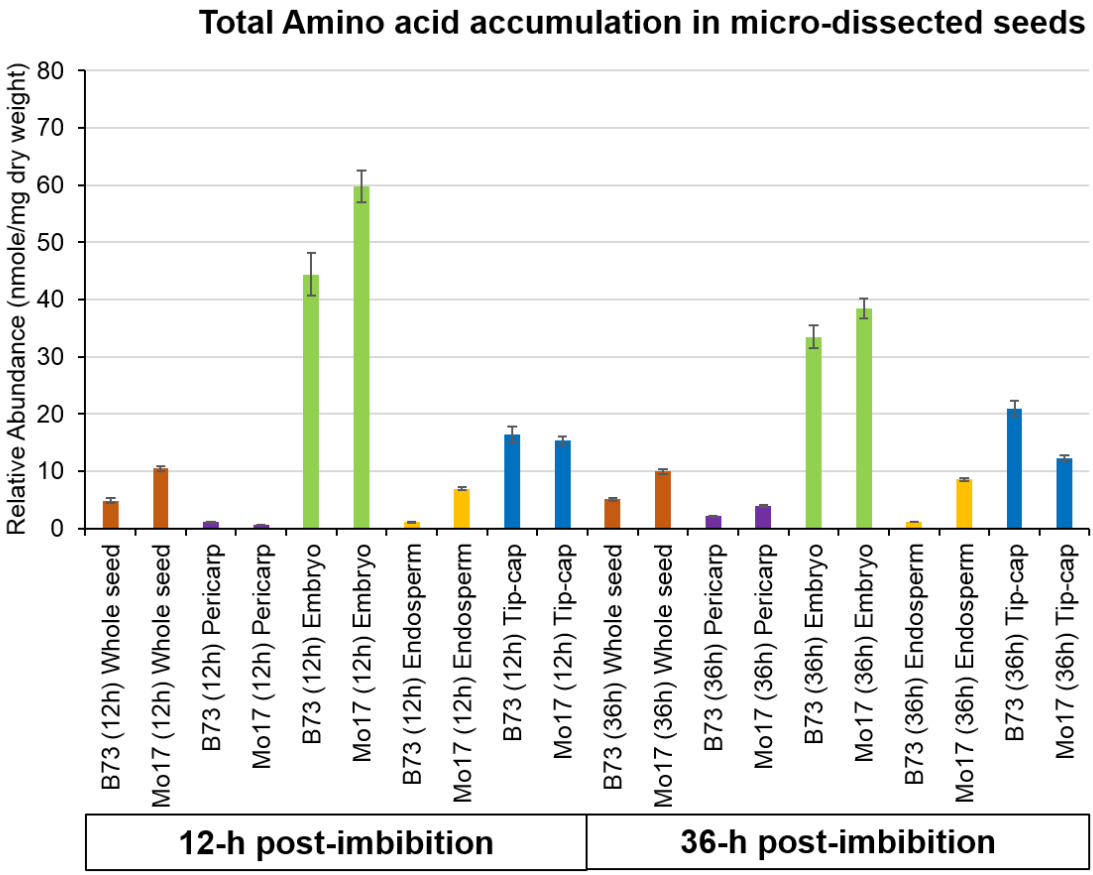


Fig. S3. Compartmentalization of amino acids in maize seeds. The y-axis denotes the relative abundance of amino acid analyzed by GC-MS from specific tissues of micro-dissected seed organs, at 12 h and 36 h post-imbibition. The calculation of relative abundance and standard error was determined from triplicate samples as described in the Material and Methods.

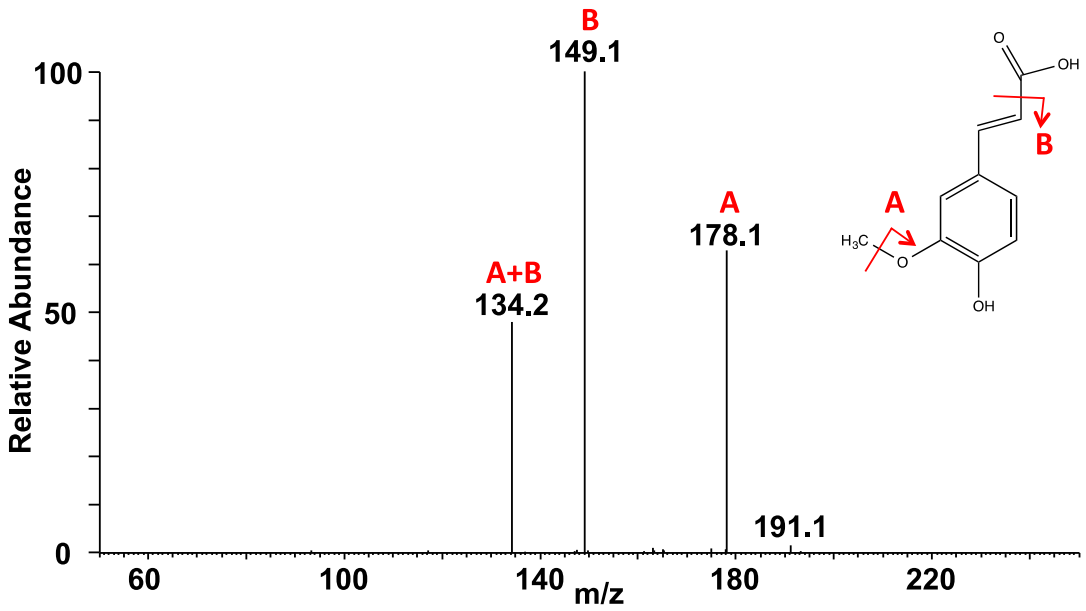


Fig. S4. MS/MS spectrum for *m/z* 193.051 from the aleurone layer of a 24 h germinated B73 seed in negative ion mode with DAN as matrix. The MS/MS fragmentation matches with the structure that of ferulic acid standard

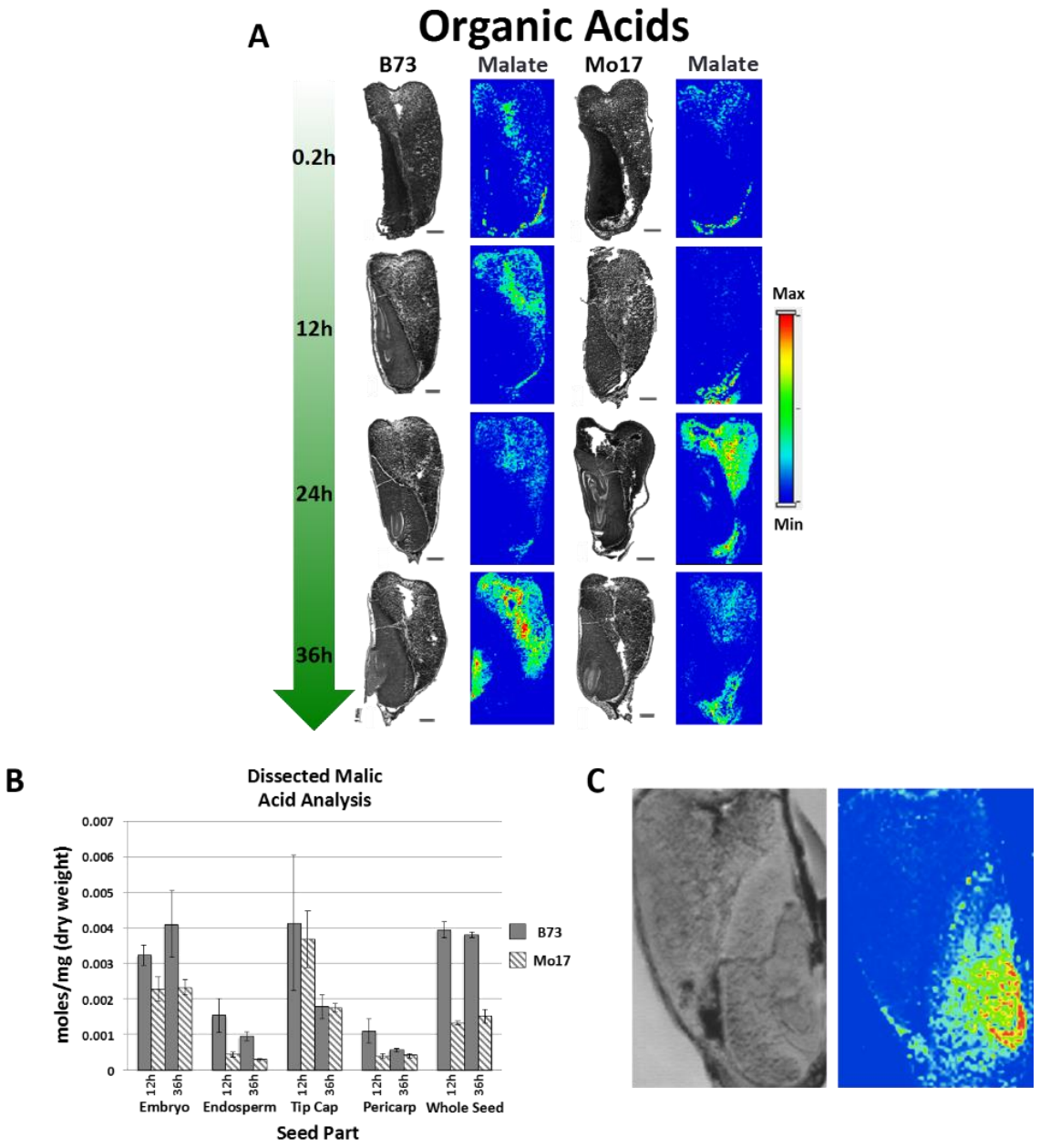


Fig. S5. Malic acid analysis. A) Images of malate distribution across all four time points in seeds on B73 and Mo17 inbreds. Images were acquired in negative ion mode with 9AA as matrix. All images are normalized to the TIC. Bar = 1mm. B) GC-MS analysis results of dissected seeds of B73 and Mo17 at 12 and 36 hours of germination. These results indicate that malic acid should be more abundant in the embryo than endosperm of the seeds, which is in contrast with the images in part A but agreement with C. C) Image of malate on a 24 hr Mo17 tissue sectioned and imaged the same day as it was harvested. The image displays malate localized to the embryo of the seed. The image was acquired in negative ion mode with 9AA as matrix and was normalized to the TIC

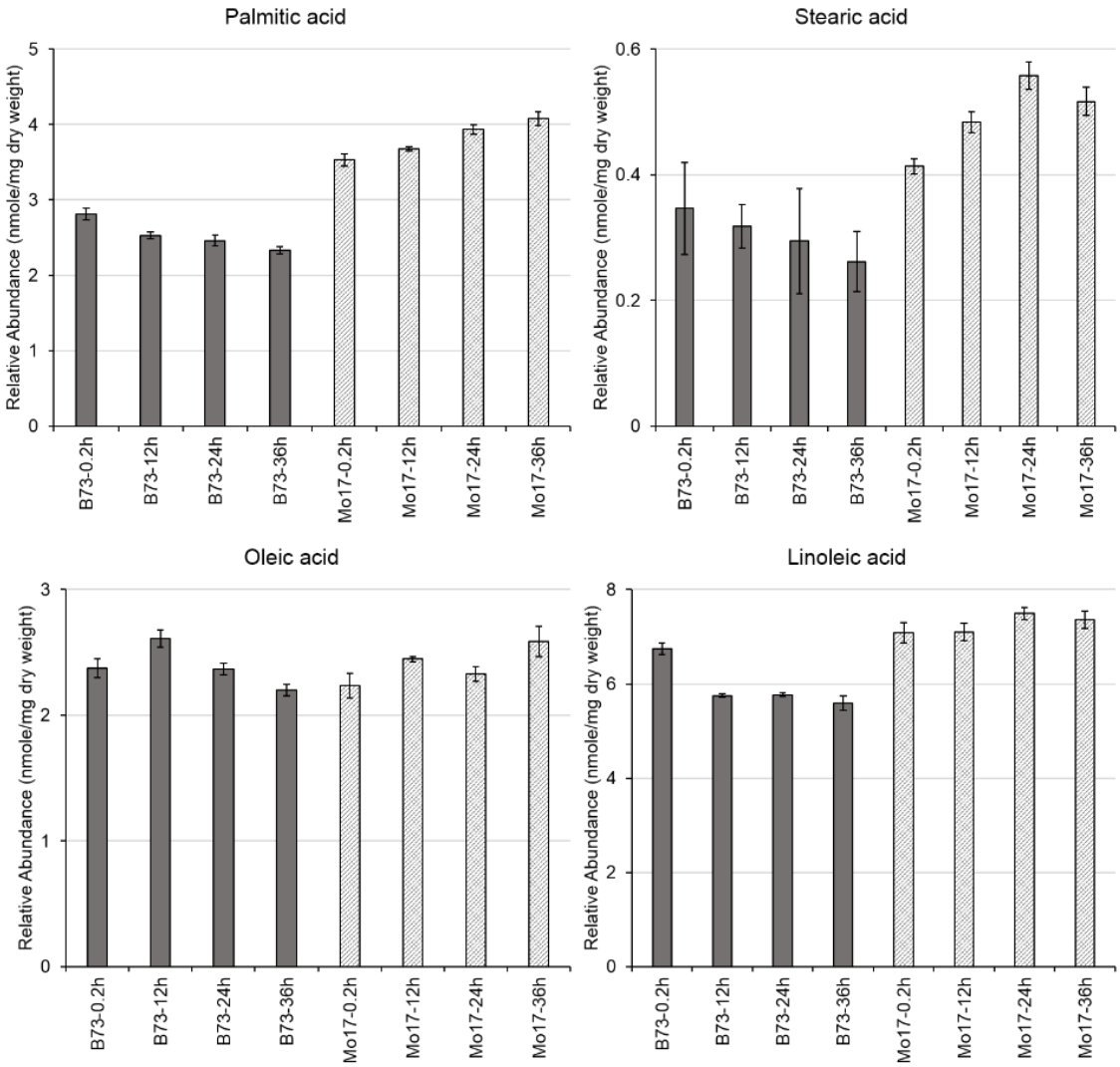


Fig. S6. GC-MS analysis of whole seeds for fatty acid species seen in MSI. From this data, only minor abundance differences between inbreds can be seen for the fatty acids. However, examination of MSI images shows significant differences in their localizations within the embryo

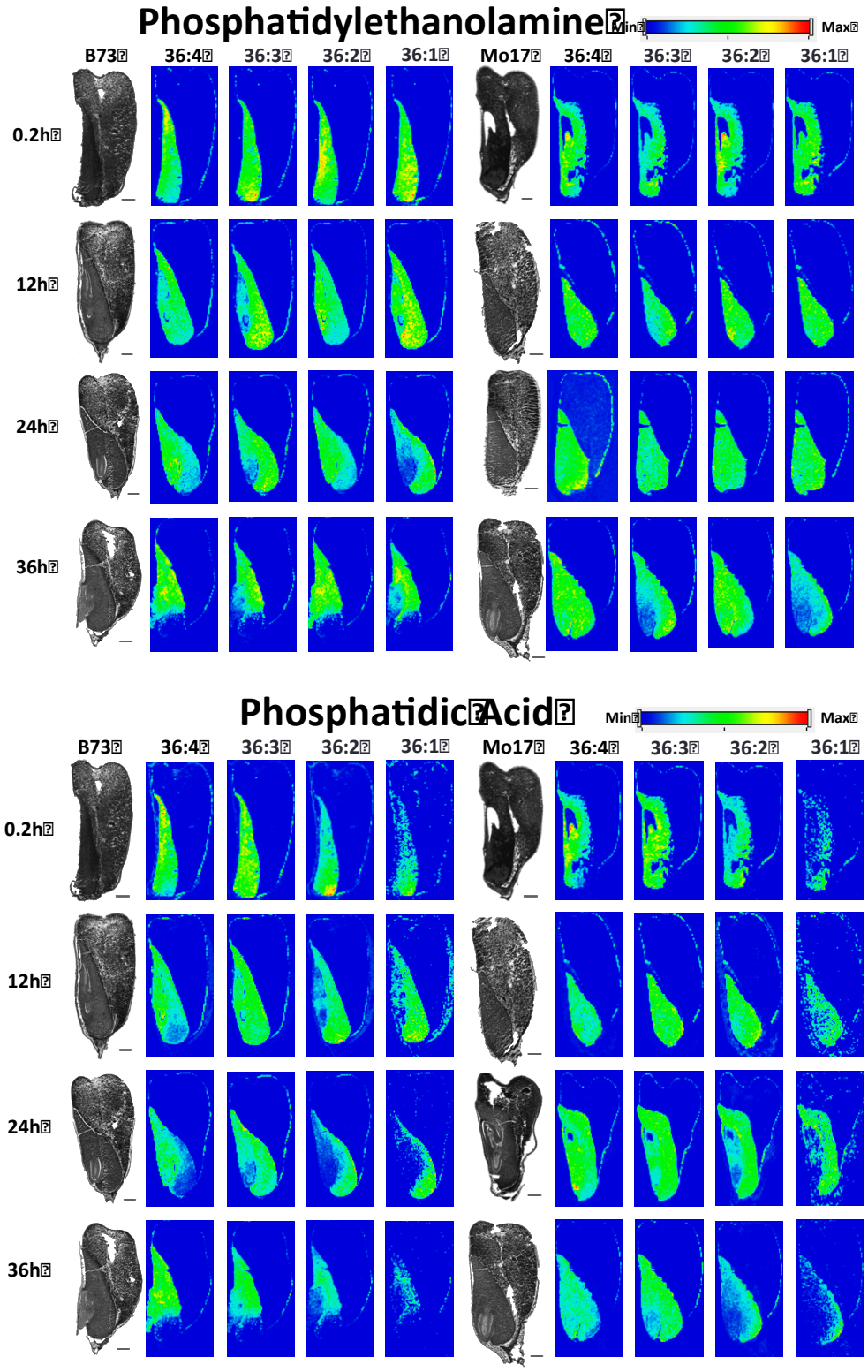


Fig. S7. Time course images for PE, PA, PI, and PC species. All imaging conditions were the same as described in Figure 5A. Scale bar: 1 mm.

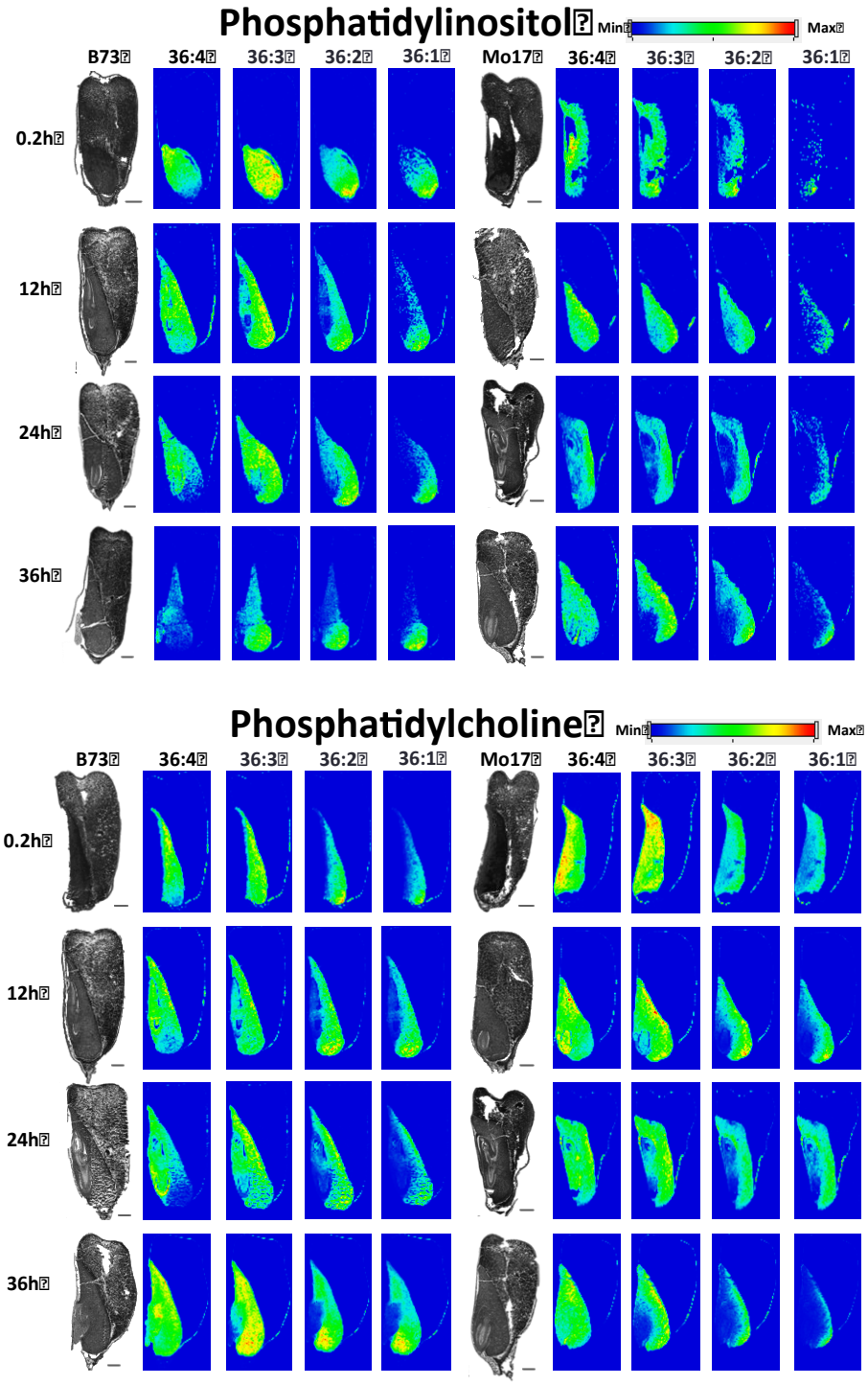

Phosphatidylinositol
Min
Max
B73
36:4
36:3
36:2
36:1
Mo17
36:4
36:3
36:2
36:1
0.2h
12h
24h
36h
Phosphatidylcholine
Min
Max
B73
36:4
36:3
36:2
36:1
Mo17
36:4
36:3
36:2
36:1
0.2h
12h
24h
36h

Fig.S7. (cont'd)

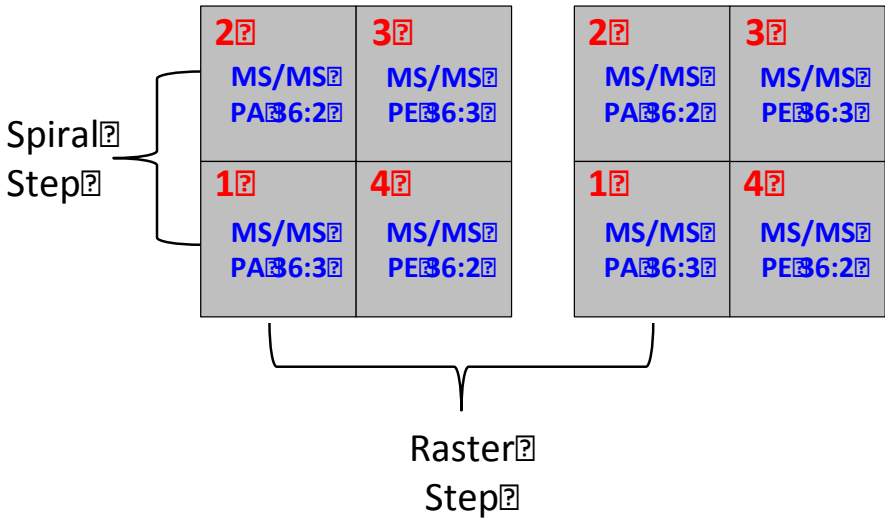


Fig. S8. Four-step imaging scheme for MS/MS of lipid species. In this setup, an individual raster step is broken up into four spiral steps, and each spiral step is assigned for MS/MS of PA 36:3, PA 36:2, PE 36:3, and PE 36:2 species, respectively. Spectra were collected in negative mode with DAN as matrix.

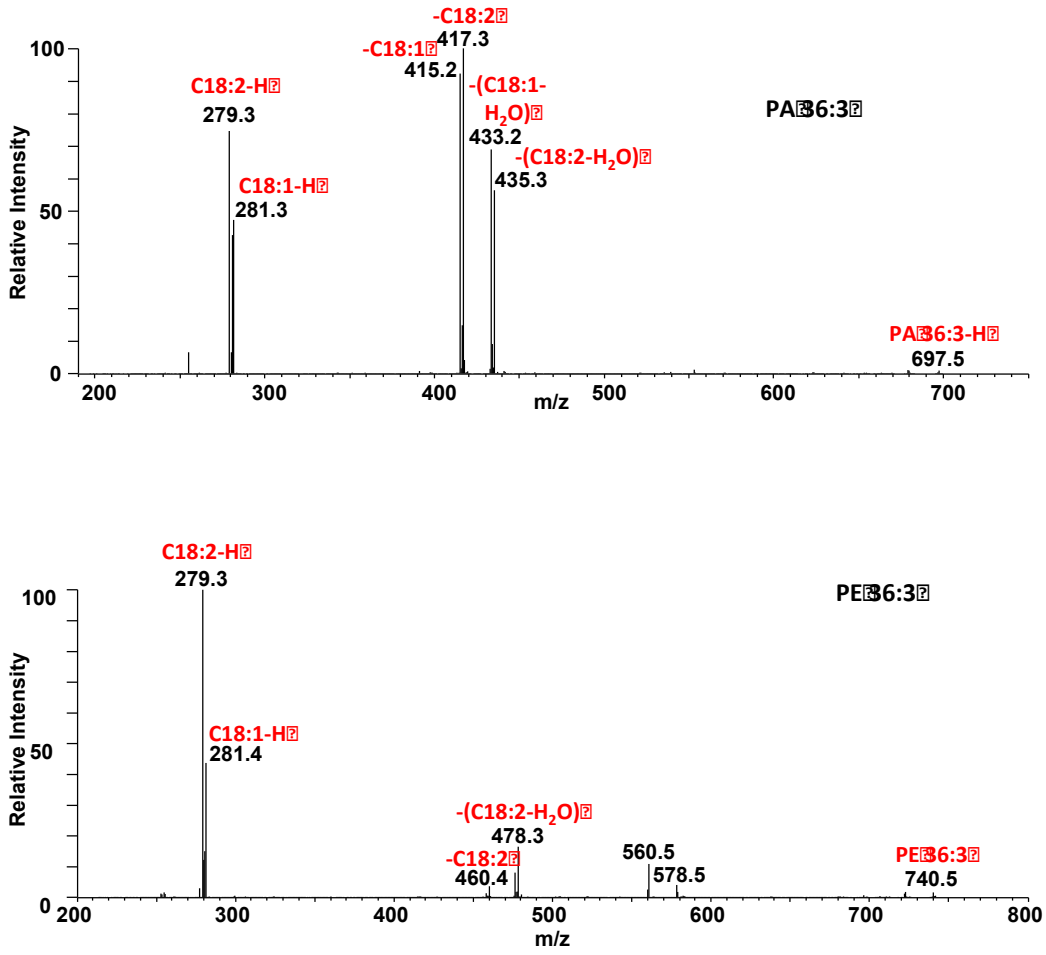


Fig. S9. MS/MS spectra from Mo17 maize seed for PA 36:3, and PE 36:3 species. Spectra were averaged over the radicle area of the embryonic axis. The spectra show that both PA 36:3 and PE 36:3 are predominantly composed of 18:1 and 18:2 fatty acid species. Spectra were acquired in negative ion mode with DAN as

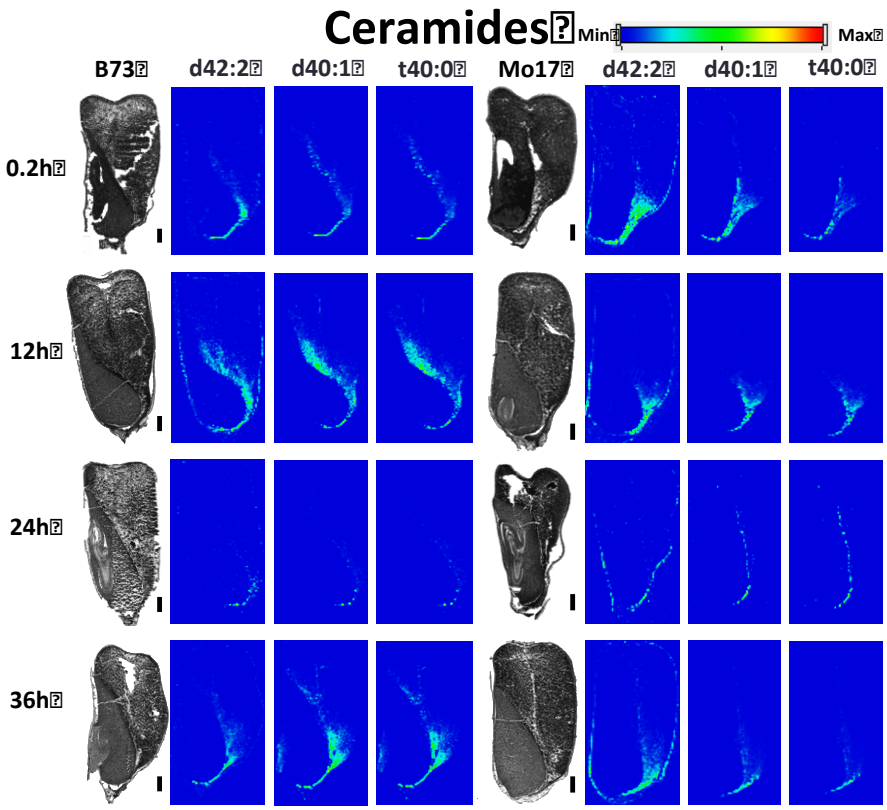


Fig. S10. Images for remaining ceramide species from Supplemental Table 1 and not shown in Figure 9. All imaging conditions are as described in Figure 9.

Table S1. Parameters used to generate mass spectrometry images, including *m/z* values, mass tolerances, and maximum and minimum scales used.

Assignment	m/z values	Tolerance[a]	Minimum and Maximum Intensity Values for Images[b]	
			B73	Mo17
Polysaccharides				
Hex_9-H_2O+K	*m/z* 1497.438	0.088	0.2h: 1.07E5 12h: 5.03E4 24h: 5.26E4 36h: 5.19E4	0.2h: 1.14E5 12h: 5.86E4 24h: 5.99E4 36h: 4.61E4
Hex_8-H_2O+K	*m/z* 1335.38	0.079	0.2h: 1.11E5 12h: 5.86E4 24h: 5.57E4 36h: 7.92E4	0.2h: 1.21E5 12h: 7.09E4 24h: 5.88E4 36h: 4.14E4
Hex_7-H_2O+K	*m/z* 1173.333	0.062	0.2h: 9.18E4 12h: 5.40E4 24h: 6.07E4 36h: 5.18E4	0.2h: 9.91E4 12h: 5.45E4 24h: 4.53E4 36h: 4.24E4
Hex_6-H_2O+K	*m/z* 1011.280	0.048	0.2h: 1.13E5 12h: 4.23E4 24h: 5.75E4 36h: 6.43E4	0.2h: 8.14E4 12h: 6.010E4 24h: 4.36E4 36h: 3.56E4
Hex_5-H_2O+K	*m/z* 849.227	0.037	0.2h: 6.20E4 12h: 2.49E4 24h: 3.46E4 36h: 4.66E4	0.2h: 6.14E4 12h: 4.63E4 24h: 3.95E4 36h: 3.10E4
Hex_2+K	*m/z* 381.080	0.012	4.00E-3	
Triacylglycerols				
TAG 54:6+K	*m/z* 917.700	0.042	0.2h: 4.36E6 12h: 1.26E7 24h: 2.92E6 36h: 1.29E7	0.2h: 8.66E6 12h: 1.62E7 24h: 5.97E6 36h: 1.08E7
TAG 54:5+K	*m/z* 919.715	0.042	0.2h: 6.99E6 12h: 1.94E7 24h: 5.09E6 36h: 1.45E7	0.2h: 8.17E6 12h: 1.77E7 24h: 4.16E6 36h: 1.03E7
TAG 54:4+K	*m/z* 921.731	0.042	0.2h: 4.85E6 12h: 1.65E7 24h: 4.18E6 36h: 9.13E6	0.2h: 6.90E6 12h: 8.27E6 24h: 1.42E6 36h: 5.32E6
TAG 54:3+K	*m/z* 923.746	0.042	0.2h: 1.88E6 12h : 7.26E6 24h: 1.56E6 36h: 4.44E6	0.2h: 2.70E6 12h: 2.36E6 24h: 4.72E5 36h: 1.31E6

Table S1. (Cont'd)				
Amino Acids				
Lysine+H	*m/z* 147.113	0.002	0.2h: 2.698E4 12h: 2.542E4 24h: 2.941E4 36h: 4.453E4	0.2h: 2.884E4 12h: 1.452E5 24h: 6.652E4 36h: 1.907E5
Arginine+H	*m/z* 175.119	0.004	0.2h: 7.576E4 12h: 1.046E5 24h: 1.507E5 36h: 6.889E5	0.2h: 1.049E5 12h: 6.844E5 24h: 8.764E5 36h: 1.469E6
Proline+H	*m/z* 116.070	0.002	0.2h: 1.344E6 12h: 3.681E5 24h: 9.704E5 36h: 3.181E6	0.2h: 3.779E4/1.926E5 12h: 7.789E5 24h: 3.364E4/2.044E5 36h: 6.840E5
Pericarp/Aleurone Compounds				
Unknown	*m/z* 135.046	0.002	0.2h: 6.02E5 12h: 7.31E6 24h: 1.02E6 36h: 4.92E6	0.2h: 6.37E6 12h: 1.07E7 24h: 1.44E7 36h: 1.19E7
Ferulate	*m/z* 193.051	0.004	0.2h: 9.11E5 12h: 4.04E6 24h: 1.47E6 36h: 3.88E6	0.2h: 4.54E6 12h: 7.45E6 24h: 6.44E6 36h: 7.73E6
Unknown	*m/z* 161.025	0.003	0.2h: 1.78E5 12h: 2.09E6 24h: 2.94E5 36h: 8.72E5	0.2h: 2.72E6 12h: 5.88E6 24h: 7.88E6 36h: 3.43E6
Combination Image Fig. 4B & C				
Malic Acid	*m/z* 133.015	0.002	4.577E2	Malic Acid
PI 34:2	*m/z* 833.521	0.036	2.586E2	PI 34:2
Ferulic Acid	*m/z* 193.051	0.004	3.751E2	Ferulic Acid
Energy Related Metabolites				
Hexose phosphate -H_2O-H	*m/z* 241.012	0.005	9.00E-3	
Glycerol phosphate -H_2O-H	*m/z* 152.996	0.003	1.50E-2	
Citrate/Isocitrate-H	*m/z* 191.020	0.004	1.00E-3	
Fatty Acids				
Palmitate (16:0)-H	*m/z* 255.233	0.005	8.00E-2	
Linoleate (18:2)-H	*m/z* 279.233	0.005	1.00E-1	
Oleate (18:1)-H	*m/z* 281.249	0.005	8.00E-2	
Stearate (18:0)-H	*m/z* 283.264	0.005	6.00E-3	
Phospholipids				
PE 36:4-H	*m/z* 738.508	0.030	0.2h: 6.60E5 12h: 4.03E6 24h: 7.67E5 36h (Supp): 2.82E6 36h (Main): 3.99E6	0.2h: 1.95E6 12h: 2.07E6 24h: 2.16E6 36h (Supp): 2.85E6 36h (Main): 2.85E6

Table S1. (Cont'd)

PE 36:3-H	*m/z* 740.524	0.030	0.2h: 4.87E5 12h: 2.410E6 24h: 4.59E5 36h (Supp): 1.52E6 36h (Main): 2.45E6	0.2h: 1.76E6 12h: 1.41E6 24h: 9.51E5 36h (Supp): 1.49E6 36h (Main): 1.49E6
PE 36:2-H	*m/z* 742.539	0.030	0.2h: 9.89E5 12h: 5.06E6 24h: 1.21E6 36h (Supp): 3.62E6 36h (Main): 4.14E6	0.2h: 3.69E6 12h: 4.45E6 24h: 2.67E6 36h (Supp): 5.71E6 36h (Main): 5.71E6
PE 36:1-H	*m/z* 744.555	0.030	0.2h: 4.75E5 12h: 1.83E6 24h: 5.83E5 36h (Supp): 1.37E6 36h (Main): 1.89E6	0.2h: 1.10E6 12h: 1.62E6 24h: 8.93E5 36h (Supp): 1.94E6 36h (Main): 1.94E6
PA 36:4-H	*m/z* 695.465	0.028	0.2h: 5.97E5 12h: 4.29E6 24h: 1.35E6 36h (Supp): 3.34E6 36h (Main): 1.25E7	0.2h: 2.62E6 12h: 4.42E6 24h: 4.85E6 36h (Supp): 6.80E6 36h (Main): 6.80E6
PA 36:3-H	*m/z* 697.481	0.028	0.2h: 8.56E5 12h: 5.70E6 24h: 1.08E6 36h (Supp): 4.18E6 36h (Main): 6.43E6	0.2h: 3.06E6 12h: 4.86E6 24h M: 5.71E6 36h (Supp): 5.15E6 36h (Main): 5.15E6
PA 36:2-H	*m/z* 699.497	0.028	0.2h: 6.51E5 12h: 3.67E6 24h: 8.34E5 36h (Supp): 2.04E6 36h (Main): 6.278E6	0.2h: 1.98E6 12h: 1.92E6 24h: 2.34E6 36h (Supp): 2.75E6 36h (Main): 2.75E6
PA 36:1-H	*m/z* 701.513	0.028	0.2h: 9.49E4 12h: 3.41E5 24h: 1.10E5 36h (Supp): 2.95E5 36h (Main): 5.46E5	0.2h: 2.85E5 12h: 2.91E5 24h: 2.79E5 36h (Supp): 4.87E5 36h (Main): 4.87E5
PI 36:4-H	*m/z* 857.519	0.039	0.2h: 2.77E5 12h: 7.75E5 24h: 1.48E5 36h (Supp): 1.18E6 36h (Main): 1.18E6	0.2h: 6.36E5 12h: 8.24E5 24h: 1.16E6 36h (Supp): 5.50E5 36h (Main): 5.50E5
PI 36:3-H	*m/z* 859.534	0.039	0.2h: 3.96E5 12h: 1.22E6 24h: 2.22E5 36h (Supp): 9.98E5 36h (Main): 9.98E5	0.2h: 7.59E5 12h: 1.06E6 24h: 1.28E6 36h (Supp): 6.30E5 36h (Main): 6.30E5

Table S1. (Cont'd)				
PI 36:2-H	*m/z* 861.550	0.039	0.2h: 6.09E5 12h: 1.43E6 24h: 3.39E5 36h (Supp): 1.86E6 36h (Main): 1.86E6	0.2h: 6.62E5 12h: 1.03E6 24h: 1.01E6 36h (Supp): 1.27E6 36h (Main): 1.27E6
PI 36:1-H	*m/z* 863.566	0.039	0.2h: 1.27E5 12h: 4.18E5 24h: 1.93E5 36h (Supp): 6.88E5 36h (Main): 6.88E5	0.2h: 1.74E5 12h: 2.53E5 24h: 3.75E5 36h (Supp): 6.33E5 36h (Main): 6.33E5
PC 36:4+K	*m/z* 820.525	0.036	0.2h: 2.15E7 12h: 2.29E7 24h: 2.36E7 36h (Supp): 3.77E7 36h (Main): 3.22E7	0.2h: 2.52E7 12h: 3.17E7 24h: 1.22E7 36h (Supp): 3.37E7 36h (Main): 3.37E7
PC 36:3+K	*m/z* 822.541	0.036	0.2h: 2.85E7 12h: 3.68E7 24h: 2.63E7 36h (Supp): 3.31E7 36h (Main): 2.48E7	0.2h: 2.83E7 12h: 3.60E7 24h: 1.11E7 36h (Supp): 2.47E7 36h (Main): 2.47E7
PC 36:2+K	*m/z* 824.557	0.036	0.2h: 2.43E7 12h: 2.10E7 24h: 2.47E7 36h (Supp): 1.80E7 36h (Main): 2.49E7	0.2h: 2.69E7 12h: 2.37E7 24h: 5.23E6 36h (Supp): 1.68E7 36h (Main): 1.68E7
PC 36:1+K	*m/z* 826.572	0.036	0.2h: 3.55E6 12h: 2.57E6 24h: 5.21E6 36h (Supp): 1.96E6 36h (Main): 4.25E6	0.2h: 3.94E6 12h: 3.48E6 24h: 5.78E5 36h (Supp): 4.39E6 36h (Main): 4.39E6
PC 34:2+K	*m/z* 796.527	0.035	2.267E7	
Ceramides				
Cer d40:1+H	*m/z* 622.614	0.023	0.2h: 9.29E5 12h: 6.83E5 24h: 6.23E5 36h: 5.68E5	0.2h: 6.83E5 12h: 8.12E5 24h: 1.67E5 36h: 1.39E6
Cer t40:0+H	*m/z* 640.625	0.024	0.2h: 8.46E5 12h: 5.89E5 24h: 5.71E5 36h: 5.59E5	0.2h: 6.69E5 12h: 7.17E5 24h: 1.39E5 36h: 1.37E6
Cer d42:1+H	*m/z* 650.645	0.026	0.2h: 1.21E6 12h: 1.22E6 24h: 3.78E5 36h: 1.57E6	0.2h: 6.77E5 12h: 1.03E6 24h: 2.85E5 36h: 1.24E6

Table S1. (Cont'd)				
Cer d42:2+H	*m/z* 664.625	0.027	0.2h: 5.53E5 12h: 3.96E5 24h: 2.51E5 36h: 8.24E5	0.2h: 3.02E5 12h: 6.08E5 24h: 1.91E5 36h: 6.25E5
Cer t42:0+H	*m/z* 668.655	0.027	0.2h: 1.34E6 12h: 9.78E5 24h: 3.39E5 36h: 1.46E6	0.2h: 4.66E5 12h: 7.19E5 24h: 2.74E5 36h: 7.61E5
Supplementary Figure 9A				
Malic Acid	*m/z* 133.015	0.002	1.00E-2	
Supplemental Figure 9C				
Malic Acid	*m/z* 133.015	0.002	1.00E-2	

a: Mass tolerance is selected considering the mass resolution at a given mass.

b: For the images normalized to TIC, the same maximum scale is used for all the images regardless of genotype or germination time point. For non-normalized images, maximum color scale is arbitrarily adjusted for each image to obtain the best contrast displaying localization. Minimum value is set to zero for the compounds normalized to TIC and one for non-normalized images unless two values are given, in which case the first is the minimum and second is the maximum.

Tables S2, S3, S4: Available online on the PMR database.

Table S5

Normalized expression level of six individual *FAD2* encoding mRNAs in radicle tissues relative to scutellar tissues of germinating seeds from inbred B73 and Mo17. The relative amount of each FAD2 transcript was normalized with the ubiquitin mRNA (GenBank Accession Number: BT018032), given by $2^{-(\Delta\Delta C_T)}$.

***FAD2* Gene ID**	***FAD2* transcript levels in radicle tissue relative to expression in scutellar tissue (B73)**			
	Time post-imbibition (h)			
	12	24	36	48
GRMZM2G056252	1.3 ± 0.4	0.9 ± 0.3	7.1 ± 0.2	4.8 ± 0.5
GRMZM2G064701	0.6 ± 0.2	0.4 ± 0.2	4.6 ± 0.2	1.4 ± 0.3
GRMZM2G169240	undetected	undetected	1.1 ± 0.2	1.4 ± 0.2
GRMZM2G169261	undetected	undetected	undetected	2.4 ± 0.2
GRMZM2G174766	undetected	undetected	undetected	undetected
GRMZM2G161792	undetected	undetected	undetected	undetected
***FAD2* Gene ID**	***FAD2* transcript levels in radicle tissue relative to expression in scutellar tissue (Mo17)**			
	Time post-imbibition (h)			
	12	24	36	48
GRMZM2G056252	1.1 ± 0.2	1.5 ± 0.2	3.2 ± 0.5	1.7 ± 0.3
GRMZM2G064701	0.2 ± 0.1	0.4 ± 0.2	6.8 ± 0.2	2.5 ± 0.2
GRMZM2G169240	undetected	undetected	undetected	46.4 ± 0.1
GRMZM2G169261	undetected	undetected	undetected	46.3 ± 0.1
GRMZM2G174766	undetected	undetected	undetected	undetected
GRMZM2G161792	undetected	undetected	undetected	undetected

Table S6

List of primer sequences used for qRT-PCR analysis

Gene ID	Forward Primer	Reverse Primer
GRMZM2G056252	5’-TACTACCACTTCGACCCGAC-3’	5’-TATGGTAGCGTCCTCTCAGCTC-3’
GRMZM2G064701	5’-TTCCTCCTCCCTGCAAATCCTG-3’	5’-CTGCTCCTGCTCATGCTTCT-3’
GRMZM2G169240	5’-CAAGCAGTGCATCTACGTTGAGC-3’	5’-GATAGAAGCAACGAGGAAGGCCA-3’
GRMZM2G169261	5’-ACATCGCCGACACGCATATC-3’	5’-GCTGGACGTAGATGCACTCT-3’
GRMZM2G161792	5’-GCAACTTGCAAAATGCGTGC-3’	5’-CGGGATGGCCTTCTTGATCT-3’
GRMZM2G174766 (inbred Mo17)	5’-AGTACTACCGCTTCGACGACA-3’	5’-TGCTTATTGCTGGCGAACCA-3’
GRMZM2G174766 (inbred B73)	5’-CAAGGAGTGCTTGTACGTCG-3’	5’-TGTAGTGTACGCCACCGTGA-3’
UBQ (BT018032)	5’-CTGGTGCCCTCTCCATATGG-3’	5’-CAACACTGACACGACTCATGACA -3’

CHAPTER 3. HIGH SPATIAL RESOLUTION MASS SPECTROMETRY IMAGING REVEALS THE GENETICALLY PROGRAMMED, DEVELOPMENTAL MODIFICATION OF THE DISTRIBUTION OF THYLAKOID MEMBRANE LIPIDS AMONG INDIVIDUAL CELLS OF MAIZE LEAF

[1]Modified from paper published in The Plant Journal (2017) 89 (4): 825-838

Maria Emilia Dueñas[1, 2], Adam T. Klein[1, 2], Liza E. Alexander[2, 3,5], Marna D. Yandeau-Nelson[4, 5], Basil J. Nikolau[2, 3,5], Young Jin Lee[1, 2]

Abstract

Metabolism in plants is compartmentalized among different tissues, cells and subcellular organelles. Mass spectrometry imaging (MSI) with matrix-assisted laser desorption ionization (MALDI) has recently advanced to allow for the visualization of metabolites at single-cell resolution. Here we applied 5- and 10-μm high spatial resolution MALDI-MSI to the asymmetric Kranz anatomy of *Zea mays* (maize) leaves to study the differential localization of two major anionic lipids in thylakoid membranes, sulfoquinovosyldiacylglycerols (SQDG) and phosphatidylglycerols (PG). The quantification and localization of SQDG and PG molecular species, among mesophyll (M) and bundle sheath (BS) cells, are compared across the leaf developmental gradient from four maize genotypes (the inbreds B73 and Mo17, and the reciprocal

[1]Department of Chemistry, Iowa State University, Ames, IA, USA
[2]Ames Laboratory-US DOE, Ames, IA, USA
[3]Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA, USA
[4]Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA, USA
[5]Center for Metabolic Biology, Iowa State University, Ames, IA, USA

hybrids B73xMo17 and Mo17xB73). SQDG species are uniformly distributed in both photosynthetic cell types, regardless of leaf development or genotype; however, PG shows photosynthetic cell-specific differential localization depending on the genotype and the fatty acyl chain constituent. Overall, 16:1-containing PGs primarily contribute to the thylakoid membranes of M cells whereas BS chloroplasts are mostly composed of 16:0-containing PGs. Furthermore, PG 32:0 shows genotype-specific differences in cellular distribution, with preferential localization in BS cells for B73, but more uniform distribution between BS and M cells in Mo17. Maternal inheritance is exhibited within the hybrids such that the localization of PG 32:0 in B73xMo17 is similar to the distribution in the B73 parental inbred, whereas that of Mo17xB73 resembles the Mo17 parent. This study demonstrates the power of MALDI-MSI to reveal unprecedented insights on metabolic outcomes in multicellular organisms at single-cell resolution.

Introduction

Thylakoid membranes play an important organizational role in maintaining the structural integrity of photosystem complexes (Quinn and Williams 1983). The thylakoid membrane is the universal site of the photochemical and electron transport reactions of oxygenic photosynthesis in cyanobacteria and plant chloroplasts. The major lipids within the thylakoid membrane are the neutral galactolipids, monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG), which together account for approximately 80% of total thylakoid lipids in plant chloroplasts (Kobayashi *et al.* 2009). The remaining thylakoid membrane lipids are mainly composed of anionic lipids sulfoquinovosyldiacylglycerol (SQDG) and phosphatidylglycerol (PG) (Demé *et al.* 2014, Sakurai *et al.* 2006); these negatively charged lipids have been shown to be responsible for the structural and/or functional integrity of photosystems I and/or II (Guskov *et al.* 2009, Jordan *et al.* 2001, Sato 2004, Umena *et al.* 2011).

A number of studies have highlighted that there are significant differences in the fatty acyl chains associated with different thylakoid lipid molecular species (Hsu *et al.* 2007). These composition patterns can be modified by both the genetic and developmental programs of the organism, and by the environmental stimuli that the organism experiences (Ben Hamed *et al.* 2005, Murata *et al.* 2003, Pál *et al.* 2007). Galactolipids (MGDG and DGDG) are characterized by an exceptionally high trienoic fatty acid content, mainly α-linolenic acid (Duchêne and Siegenthaler 2000, Nishihara *et al.* 1980), whereas the major lipid species in SQDG are enriched in palmitic acid. PG is a distinct thylakoid lipid, in that its acyl composition vastly differs depending on the plant species.

Unlike C3 plants, *Zea mays* (maize) employs a highly efficient C4-type photosynthesis that is differentially regulated by cellular compartmentalization of photosynthetic carbon assimilation in the mesophyll (M) cells, and photosynthetic carbon reduction in the bundle sheath (BS) cells, creating a two-cell metabolism system. In addition to the metabolic flux that occurs between these two cell types, ultrastructural and photosystem assays have revealed that the complexity of C4 photosynthesis is also dependent on the composition differences of the chloroplast envelopes and thylakoid membranes (Manandhar-Shrestha *et al.* 2013). One of these differences is associated with the acyl-chain composition of PG, which appears to be different in the thylakoid membranes of M and BS cells. Specifically, PG with a 16:1 fatty acid at the sn-2 position (e.g., PG 16:0/16:1) is highly enriched in M chloroplasts, and PG with 16:0 at the sn-2 position (e.g., PG 16:0/16:0) is primarily localized in BS chloroplasts (Nishihara, et al. 1980). Moreover, the acyl-chains of PG are correlated with developmental stages as well as chilling sensitivity. For example, plants with high levels of disaturated thylakoid PGs are shown to be sensitive to chilling (Roughan 1985); however, this is not without debate (Kaniuga *et al.* 1999).

In plants, grass species serve as an excellent model to study the establishment of various functions relative to the developmental gradient of leaves, in which the youngest cells are situated at the base of the leaf, and the oldest and most mature cells are situated at the leaf tip (Nelson and Langdale 1992). Cellular and photosynthetic differentiation occurs in a controlled manner in a basipetal, or tip-to-base, fashion (Evert *et al.* 1996, Kirchanski 1975, Leech *et al.* 1973) with proplastids present at the base of the leaf blade and fully differentiated BS and M chloroplasts present at the tip (Evert, et al. 1996, Majeran and van Wijk 2009, Nelson and Langdale 1992). A number of studies have characterized the transcriptomic and metabolomic changes across the developmental gradient of maize leaves (Li *et al.* 2010, Pick *et al.* 2011, Wang *et al.* 2014). Differential transcriptome (Sharpe *et al.* 2011, Tausta *et al.* 2014) and proteome (Majeran *et al.* 2005, Majeran *et al.* 2008, Manandhar-Shrestha, et al. 2013) expression have been identified between M and BS cells; however, because these omics studies were performed through the extraction of ground tissue, *in situ* localization for metabolites or proteins, for example, were not preserved.

Mass spectrometry imaging (MSI) has become a valuable tool for analyzing spatial distributions of a wide range of compounds directly on or within plant and animal tissues (Angel and Caprioli 2013, Jungmann and Heeren 2012, Lee *et al.* 2012, Sturtevant *et al.* 2016, Svatos 2010, van Hove *et al.* 2010, Zaima *et al.* 2010). Different ionization techniques have been adopted for MSI; however, matrix assisted laser desorption ionization (MALDI) is most attractive in terms of high spatial resolution, sensitivity, and chemical versatility, which is essential for *in situ* single-cell resolution imaging. The spatial resolution of MALDI-MSI has become routinely available in the range of 20-30 μm, and has been demonstrated to as high resolution as 2.5 μm (Zavalin *et al.* 2015).

Recently, we have established a MALDI-MSI technical platform with a MALDI-linear ion trap Orbitrap mass spectrometer for a high-spatial resolution of 5-10 μm and have demonstrated the visualization of the distribution of a number of different metabolites in cross-sections of maize leaves (Korte *et al.* 2015). Cell-type specific non-uniform distribution of PG, previously indicated by bulk chloroplast analysis from physically separated BS and M cells (Nishihara, et al. 1980), was confirmed using this single cell direct *in situ* MSI analysis platform (Korte, et al. 2015). In previous work, we have achieved a 5-μm spatial resolution using an oversampling method with a 9-μm laser spot size. In the present work we used an improved laser spot size of 6-7 μm with minimal oversampling which allows for a clearer image especially for small cells, and applied the high-resolution MALDI-MSI platform to explore quantitative fatty acyl distributions of PG and SQDG along the developmental gradient of maize leaves in two inbred lines, B73 and Mo17, and the reciprocal hybrid lines, B73xMo17 and Mo17xB73.

Results

Morphology and Lipid Distribution during Leaf Development

The fatty acid composition of thylakoid membrane lipids was studied at four consecutive developmental zones in the maize seedling leaf: the basal zone, maturing proximal zone, maturing midpoint zone, and mature distal zone, as defined by Li *et al.* (2010). The overall workflow is shown in Figure S1 and described in detail in the Experimental Procedures section. The third true leaf was harvested from seedlings of inbreds B73 and Mo17, and reciprocal hybrids B73xMo17, and Mo17xB73 11-13 days after imbibition, at a seedling height of 14-18 cm. Each sampled leaf was immediately flash frozen and cryosectioned at four specific positions along the developmental gradient of the leaf, as illustrated in Figure S2: (i) basal zone, ~1.5 cm below the ligule of leaf 2;

(ii) proximal, ~1.5 cm above the ligule of leaf 2; (iii) midpoint, halfway between proximal end and the leaf tip; and (iv) Distal, ~2 cm from the leaf tip.

High-resolution optical images and anatomical assignments are shown in Figure S3 for the distal and proximal sections of a B73 leaf. In the optical images of Figures 1 and 2, both inbred lines and the two hybrid lines exhibit the characteristic C4 Kranz anatomy (Brown 1975, Haberlandt 1882), in which concentric rings of BS cells encircle closely spaced veins (vascular bundles, VBs), and are surrounded by M cells. Chloroplasts are mostly located along the boundary of M and BS cells. At the basal zone of the third leaf, Kranz anatomy can be distinguished but is not fully differentiated and few mature chloroplasts are observed (Kirchanski 1975). Because growth rates differ among the four genotypes, leaves from the B73 inbred and the B73xMo17 hybrid were harvested 1-2 days prior to the Mo17 inbred and Mo17xB73 hybrid, such that leaves from each genotype were of similar lengths. Nevertheless, as revealed by the optical images, the sizes of the cells are smaller in inbred Mo17 and hybrid Mo17xB73.

Figures S4-S7 show the MS images for all the molecular species of PG and SQDG obtained at 10-μm spatial resolution, as well as several additional metabolites that are associated with photosynthesis, i.e., chlorophyll *a*, plastoquinone/plastoquinol, and carotene. Assignments are based on accurate mass measurements, and their identities are also confirmed by separate MS/MS measurements for the lipids (Figures S8 and S9). The presence and cellular distribution of these metabolites and thylakoid membrane lipids were similar in replicate imaging experiments conducted on three independent seedlings. Figures 1 and 2 present the representative images of three thylakoid membrane lipids (PG 32:0, PG 32:1, SQDG 34:3) and chlorophyll *a*, comparing between the inbreds and the reciprocal hybrids, respectively.

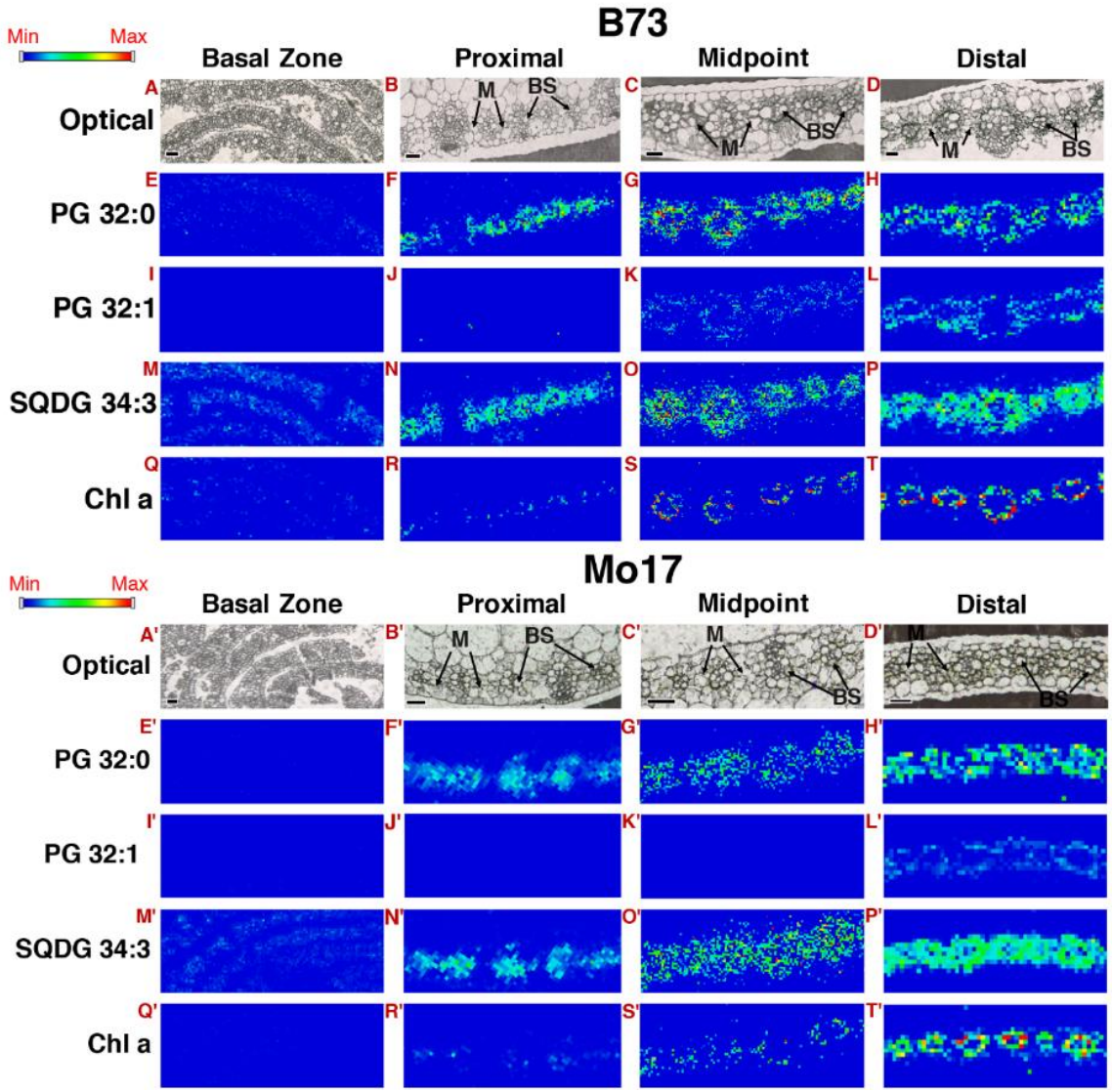


Figure 1 Ten-micron resolution MS images of representative lipids at four developmental sections of the third seedling leaf from inbreds B73 (upper panels) and Mo17 (lower panels). The MS images are of PG 32:0 (*m/z* 721.503; E–H, E′–H′), PG 32:1 (*m/z* 719.480; I–L, I′–L′), SQDG 34:3 (*m/z* 815.501; M–P, M′–P′) and chlorophyll *a* [sum of fragments at *m/z* 591.261 (pheophorbide *a*), 613.232 (chlorophyllide *a*) and 870.566 (pheophytin *a*); Q–T, Q′–T′]. The ion signals are normalized to the total ion count and the maximum value of 1.3×10^{-2} was used for all the images. Scale bars: 50 μm.

Consistent with the photosynthetic developmental gradient that is visualized by the progression from etiolation at the base to green pigmentation at the tip of the seedling leaf, the ion signal intensity for chlorophyll *a* increases from the base of the leaf to the tip. MS imaging visualization of the cross-sectional distribution of chlorophyll *a* indicates that it is highly abundant

in BS cells (Figures 1 and 2). Higher magnification MS images indicate the occurrence of chlorophyll *a* in M cells (Figures S10 and S11), but it occurs at much lower abundance than in BS cells. In contrast to chlorophyll *a* that can absorb UV at 355 nm, which is the wavelength of the laser used for these experiments and can thus be ionized even without a matrix, chlorophyll *b* does not absorb at this wavelength, and could not be detected in these experiments. The fact that

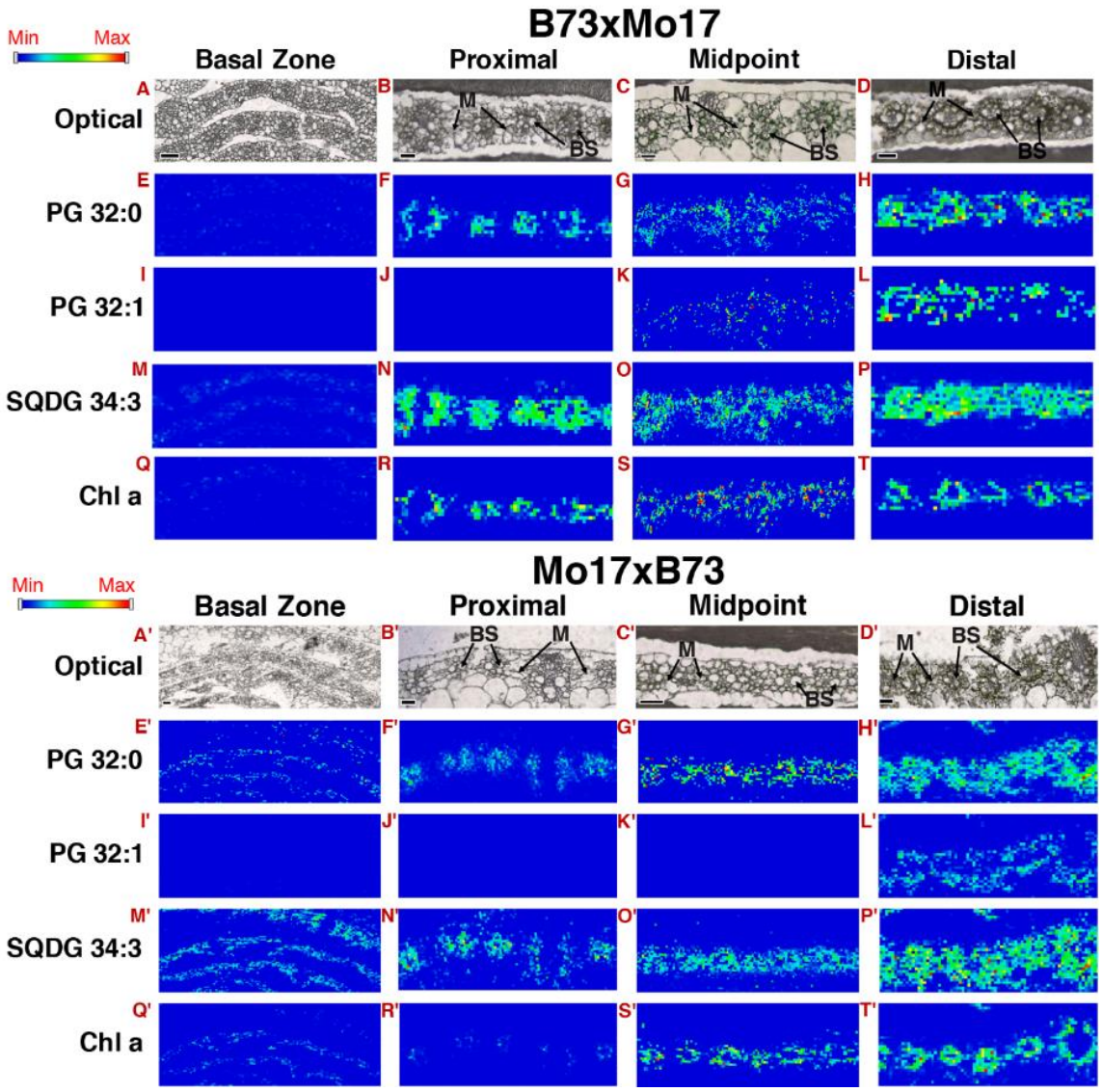


Figure 2 Ten-micron resolution MS images of representative lipids at four developmental sections of the third seedling leaf from reciprocal hybrids B73 × Mo17 (upper panels) and Mo17 × B73 (lower panels). The ion signals are normalized to the total ion count and the maximum values used were 3.0×10^{-3}, 3.0×10^{-3}, 1.3×10^{-2} and 1.3×10^{-2}, for PG 32:0, PG 32:1, SQDG 34:3 and chlorophyll *a*, respectively. Scale bars: 50 μm.

chlorophyll *a* is highly concentrated in the thylakoid membranes of BS cells can be used to guide the localization of SQDGs and PGs between BS and M cells (Kirchhoff *et al.* 2013). The chlorophyll fluorescence image obtained for a consecutive tissue section agrees well with the MS image, thereby validating our technology (Figure S12). It must be noted that chlorophyll *a* may act as a matrix itself as a result of its laser absorption, but its effect is expected to be minimum compared with the 1,5-diaminonaphthalene (DAN) matrix according to our experiment on standard samples (Figure S13).

SQDG 34:3 distribution is homogeneous between both the photosynthetic cell types, regardless of genotype and stage of development (Figure 1M-P and 1M'-P'; Figure 2M-P and 2M'-P'). In contrast, PGs are differentially distributed among M and BS cells depending on the molecular species, and this non-uniform distribution is affected by genotype, being different among the two inbreds and the hybrids. As we previously reported for the midpoint section (Korte, et al. 2015), in inbred B73, PG 32:0 is primarily located in BS cells (Figure 1E-H) and PG 32:1 is almost exclusively present in M cells (Figure 1I-L); this distribution is unaffected by leaf development, but the MS images are clear only for the midpoint and distal sections and are absent or unclear at the basal zone and the proximal end. In Mo17 however, PG 32:1 is not observed in the immature midpoint section as well as the basal and proximal zones, and accumulates to low levels in M cells only in the distal section (Figure 1I'-L'). In contrast to B73, PG 32:0 in Mo17 has a broader distribution (Figure 1E'-1H') occurring in both BS and M cells and this distribution is not affected by development. This broader distribution of PG32:0 in Mo17 across developmental stages is very similar to the observed distribution for SQDG 34:3 (Figures 1M'-1P'). In the hybrids, the distribution of these lipids generally mirrors the maternal inheritance pattern, with B73xMo17 showing similar patterns to that of the female parent, B73, and in the Mo17xB73 hybrid the pattern

is similar to the female parent, Mo17 **(**Figure 2). Specifically, PG 32:0 is located mostly in BS cells for B73xMo17 and in both M and BS cells for Mo17xB73. These distribution patterns are more evident at 5 μm spatial resolution, especially when overlaid with the optical images (Figures 4, S10, and S11).

Relative Quantification of Lipids

The relative quantification of each SQDG and PG lipid molecular species was determined by averaging the mass spectra of the MALDI-MSI dataset and then calculating the relative abundance by normalizing with the total ion signal of each lipid class. This relative quantification normalized within the same class has been shown to be a reliable approach to compare different fatty acyl chains within the same lipid class (Horn *et al.* 2012). To further evaluate the analytical variation of our approach, we performed MALDI MSI analysis of five consecutive sections of maize leaf and compared their relative quantification of SQDG and PG species (Figure S14). As shown in these data, the deviation between consecutive sections is minimal which suggests that the analytical variation in this analysis is negligible compared with biological variations. In electrospray ionization, it is known that the ionization efficiency of the lipids is predominantly affected by polar head groups and that the length of the acyl chain and the degree of saturation only minimally affects the ionization (Han and Gross 1994, Han *et al.* 2012). To demonstrate that this is also true in MALDI-MS, we compared MALDI-MS and ESI-MS for total lipid extract of maize leaf. As shown in Figure S15, they showed a similar trend, except for some minor differences in PG 34:4.

Figure 3 shows the relative abundance of the individual molecular species of SQDG and PG at the four developmental positions of the leaf. SQDG composition is similar across genotypes at each of the leaf positions that are exposed to light (i.e., the proximal, midpoint and distal

positions), with SQDG 34:3 being the predominant species (~ 60-70 mol% in each genotype), SQDG 32:0 comprising 10-20 mol%, and SQDG 34:2 and 36:6 each comprising 5-15 mol%. In the etiolated tissue of the basal zone, however, SQDG 34:2 is the primary thylakoid lipid and

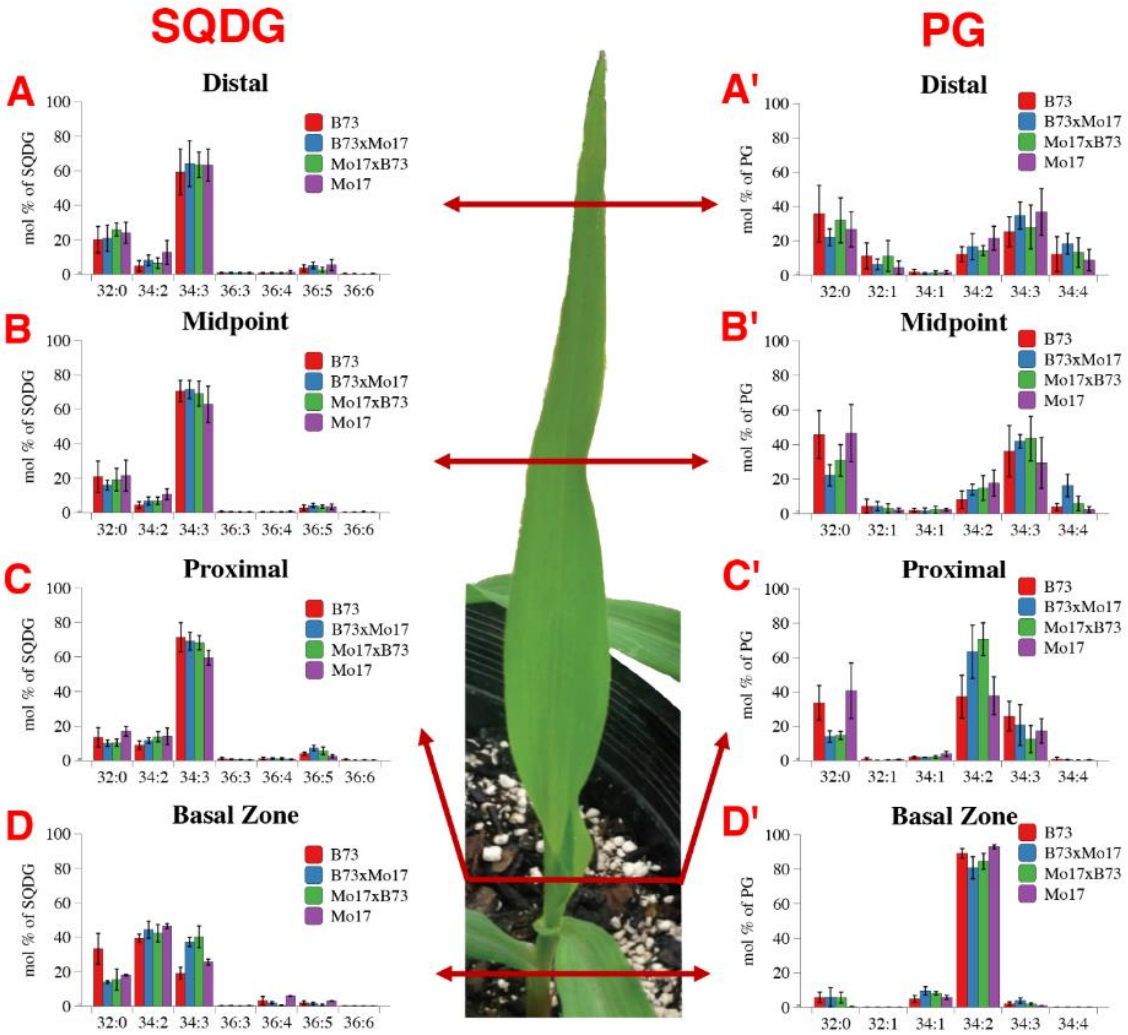


Figure 3 Quantitative comparison of SQDG (left) and PG (right) molecular species from MS images of the distal (A, A′), midpoint (B, B′), proximal (C, C′) and basal zone (D, D′) sections of maize leaves in four genotypes.

comprises 40 mol%, which is significantly higher than any of the exposed portions of the leaf that contain 5-15 mol% ($P < 0.001$ for all genotypes). SQDG 34:3 is no longer the most dominant species in the basal zone and the relative abundance is significantly less than is found in the leaf zones exposed to light ($P < 0.001$ for all genotypes). Also notable is the fact that there is no

evidence for the presence of palmitoleate (16:1 fatty acyl) containing SQDG species, such as SQDG 32:1 or 34:1. This suggests that the observed SQDG species are composed of 16:0 and 18:x fatty acids (FAs) (x = 0-3). FA 18:2 is most abundant in the basal zone of the leaf (i.e., SQDG 34:2), but is rapidly replaced by FA 18:3 (i.e., SQDG 34:3) as the leaf develops from proximal to distal portions. These observations are supported by complementary MS/MS experiments (Figure S8), and are in agreement with the previous characterization of thylakoid lipids from isolated BS and M cells (Nishihara, et al. 1980). There are some statistically supported genotypic differences in the relative abundances of these species but these are mostly minor differences.

The fatty acid compositional changes in PG are similar to SQDG, in that 18:2-containing PG is abundant at the basal zone (i.e. high abundance of PG 34:2), but is gradually replaced by 18:3-containing PG species, correlating with the emergence of the seedling leaf into direct illumination (i.e. an increase of PG 34:3 and a decrease of PG 34:2); however, accumulation patterns of PG species are more dynamic across the developmental gradient compared with SQDG species. Specifically, the decrease of 34:2 and the increase of 34:3 species are more gradual in PG as development occurs, whereas there are only minimal changes in SQDG for the leaf sections exposed to light. In PG lipids, PG 34:3 is present at trace levels in the basal zone and increases up to 20-40 mol% as the leaf develops, but is not the most predominant species in any of the developmental stages. In contrast, the analogous SQDG 34:3 is the predominant species of SQDG (~60 mol%) in all leaf zones with the exception of the etiolated basal zone, in which it comprises only ~20-40 mol%.

Phosphatidylglycerol (PG) composition also differs from SQDG composition relative to FA 16:1 and FA 16:0-containing species. Low but distinct levels of FA 16:1-containing PG species (i.e., PG 32:1, PG 34:1 and PG 34:4) accumulate as the leaf develops, whereas FA 16:1-containing

species are completely absent in SQDG. This is also consistent with MS/MS analysis, which revealed the existence of two species of PG 34:3, 18:3/16:0 and 18:2/16:1 (Figure S9C), whereas SQDG 34:3 is composed of just 18:3/16:0 (Figure S8B). PG species containing two FA 16:0 (i.e., PG 32:0) accumulate at very low levels in the basal zone, and increase to 20-35 mol% in the distal sections, in contrast to SQDG 32:0 which has no or minimal change in the relative abundance across the developmental gradient ($P > 0.05$ for most pair-wise comparison). In addition, genotypic differences in PG 32:0 accumulation are observed in the proximal sections when comparing the inbred parents with the reciprocal hybrids ($P < 0.05$), which gradually disappears in later developmental zones, with almost no difference at the distal section. The major contrasts in FA 16:0 and FA 16:1 compositions between PG and SQDG species suggest that such changes may be correlated with the maturity of the photosynthetic cells. It is important to note that PG 36:*x* species were not detected in this study. They are observed in both ESI-MS and MALDI-MS analysis of the total extract, but in very low abundance, at <0.5% (Figure S15). This is attributed to the high detection limit of the high-resolution MS imaging experiment because of the limited sampling volume.

Analysis of Cellular Distributions via High-resolution Mass Spectrometry Images

Close examination of optical microscopic images from leaf cross sections in Figures 1 and 2 indicates distinct genetics-based size differences between M and BS cells. Specifically, inbred Mo17 and hybrid Mo17xB73 BS cells are 10-20 μm in diameter, whereas BS cells in inbred B73 and hybrid B73xMo17 are twice as large at 20-40 μm in diameter. Because of the small BS cell sizes, 10-μm resolution used in Figures 1 and 2 is not sufficient to clearly distinguish the cellular distribution of some lipids in Mo17 and Mo17xB73. To clarify the cell type-specific localization of PG species, an MS imaging experiment was performed at 5-μm resolution for the midpoint and

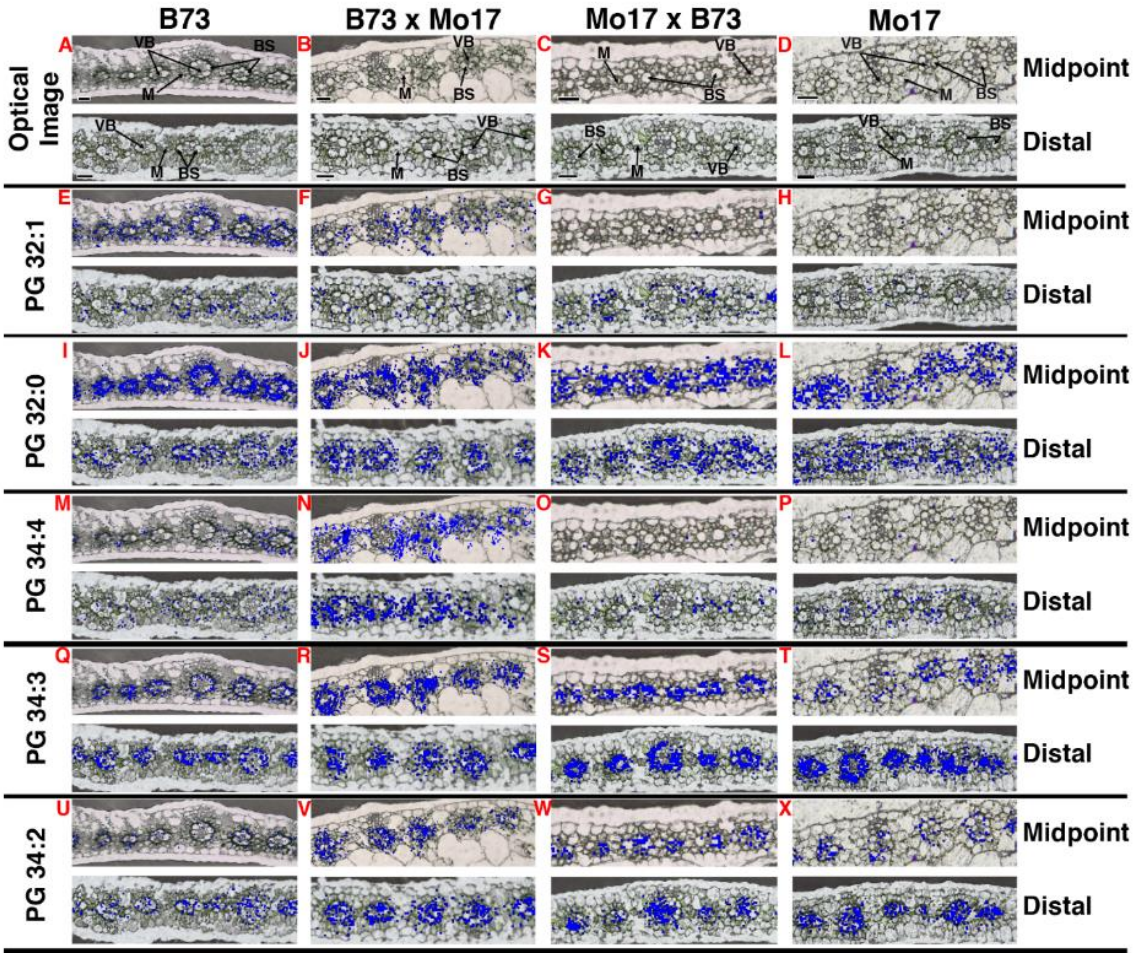


Figure 4 Five-micron resolution MS images of each PG molecular species at the midpoint and distal sections of leaf 3 from the four genotypes overlain with optical images. Scale bars: 50 μm. Note that the scale bar is bigger for Mo17 and Mo17 × B73.

distal sections. The MS images of five PG species are compared side-by-side between each genotype and between the two developmental stages (Figure 4). For better visualization of the cellular localization of these lipids, the generated MS images are overlaid and aligned with optical microscope images obtained in parallel from consecutive leaf cross sections. The MS images of PG species are also compared with those of SQDG 34:3 and chlorophyll *a* as shown in Figures S10 and S11 for the midpoint and distal sections, respectively.

For inbred B73 and hybrid B73xMo17, PG 32:1 and PG 34:4 are primarily localized in M cells in both the midpoint and distal sections zones (Figures 4E, F, M, N), whereas the other three

PGs (PG 32:0, 34:3 and 34:2) are accumulated at low levels in M cells and at much higher levels in BS cells (Figures 4I, J, Q, R, U, V). Consistent with the mol% distribution of PG 32:1 data shown in Figure 3B', PG 32:1 is almost absent in Mo17 and Mo17xB73 images at the midpoint (Figures 4G, H), and appears to be substituted by PG 32:0, which is present in both M and BS cells (Figures 4K, L). PG 32:1 is clearly visible at the distal section for Mo17xB73 (Figure 4G) and slightly for Mo17 (Figures 4H), and is localized mostly in M cells, like B73 and B73xMo17 (Figures 4E, F). PG 32:0, however, is still present in both BS and M cells for Mo17 and Mo17xB73 (Figures 4K, L). For both midpoint and distal sections, the PG 34:2 (Figures 4U, X) and PG 34:3 (Figures 4Q, T) species show similar localization patterns across all genotypes, in that both species are concentrated in BS cells. The two species that constitute PG 34:3, namely 18:2/16:1 and 18:3/16:0, cannot be distinguished here. PG 34:4 (Figures 4M-P) shows localization concentrated in the M cells; this localization pattern is similar across all genotypes in both midpoint and distal sections. Similar distributions between midpoint and distal sections for these PG species further confirm the differential localization is not changing as development occurs.

MS/MS imaging of PG 34:3 molecular species

Because the two structural isomers of PG 34:3 (i.e., 18:2/16:1 and 18:3/16:0) could not be resolved via MS imaging alone, it was unclear whether these isomers differed in their abundance or cellular localization. Therefore, to address this question, MS/MS imaging experiments were performed in 5-µm high spatial resolution to distinguish the structural isomers of PG 34:3, and to determine whether their relative quantification or distribution was affected by genotype and/or leaf development. Figure 5 compares the relative abundance of PG 18:3/16:0 and PG 18:2/16:1 species that were extracted from MS/MS imaging datasets. For each genotype, PG 18:2/16:1 occurs at minimal levels in the basal zone (~5%), and the abundance increases along the developmental

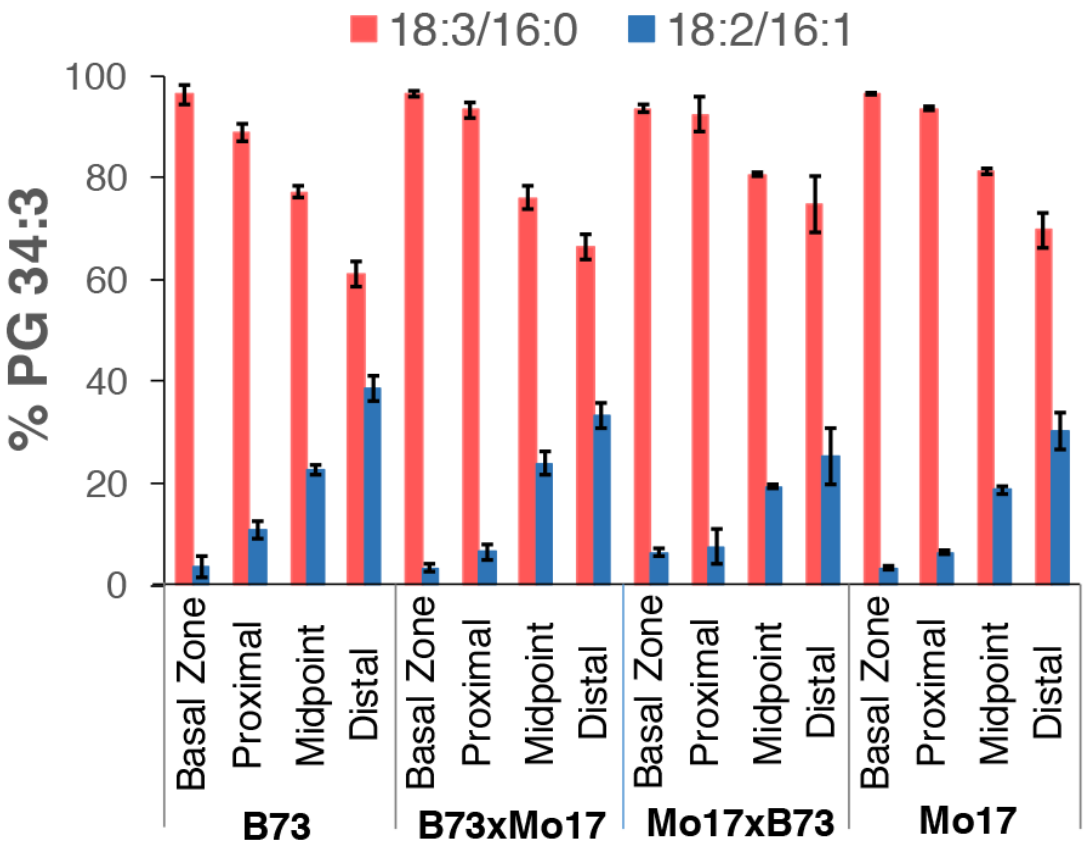


Figure 5 Quantitative comparison of the distribution of PG 18:3/16:0 and PG 18:2/16:1 as affected by leaf development among the four genotypes. These data were extracted from MS/MS imaging data sets of PG 34:3.

gradient and comprises 30-40% of PG 34:3 at the distal end. Mirroring the above change in PG 18:2/16:1, PG 18:3/16:0 accumulation decreases from ~95 to 60% across the developmental gradient of the leaf. The genotypic differences among these PG molecular species are minor: PG 18:2/16:1 reaches slightly higher levels in the distal section of B73 compared with Mo17xB73 or Mo17 ($P = 0.02$ and 0.03, respectively).

In Figure 6, 5-μm high-spatial resolution MS/MS images acquired at the midpoint zone show the localization patterns of the two individual isomers that constitute PG 34:3. Irrespective of the genotypes imaged, the cellular localization of PG 18:2/16:1 is evenly distributed between M and BS cells. In contrast, the cellular localization of PG 18:3/16:0 is mostly concentrated within BS cells. This divergence between distributions of PG 18:2/16:1 and PG 18:3/16:0 is in contrast

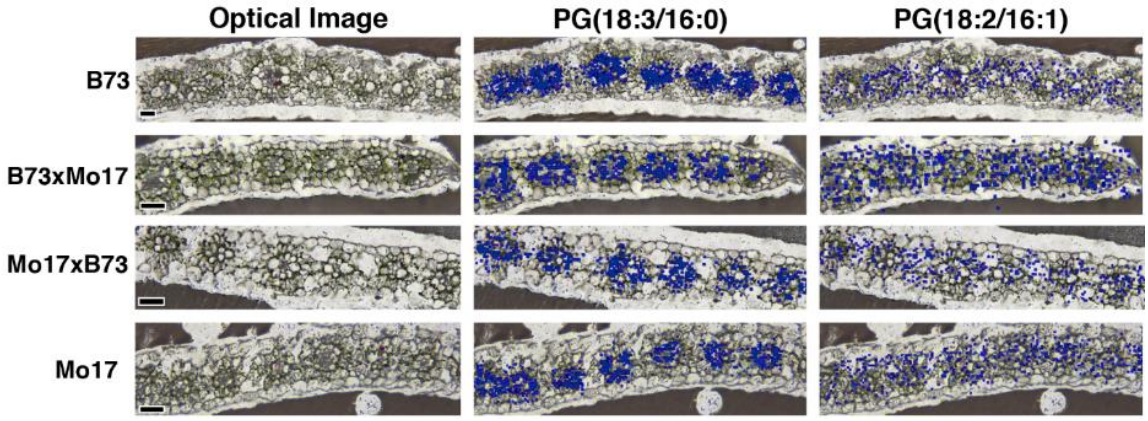


Figure 6 Five-micron resolution MS/MS images of two structural isomers of PG 34:3 at the midpoint section of maize leaves. The ion abundances for PG 18:3/16:0 is determined from the sum of $[M-16:0]^-$, $[M-(16:0-H_2O)]^-$, and $[M-(16:0+glycerol)]^-$, and PG 18:2/16:1 is determined from $[M-16:1]^-$, $[M-(16:1-H_2O)]^-$ and $[M-(16:1+glycerol)]^-$. Scale bar: 50 μm.

to the localization of PG 32:0 (16:0/16:0) and PG 32:1 (16:0/16:1). As seen in Figure 4, PG 16:0/16:1 is mostly present in M cells, whereas PG 16:0/16:0 distribution differs among genotypes with localization to mostly BS cells in B73 and B73xMo17, and both M and BS cells in Mo17 and Mo17xB73. It suggests that the occurrence of the 16 carbon fatty acid in these lipids is not the only factor that determines their localization.

Table 1 summarizes the cell-specific localization of PG molecular species for the four genotypes from Figures 4 and 6 (detailed data are presented in Table S1). Quantitative analysis was performed for the intensity ratio of these PG molecular species in unit areas of BS versus M cells, as shown in the parentheses in Table 1, which confirms the visual localization of these species. These data together show that there is a common trend for FA16:0-containing PGs (i.e. PG 16:0/16:0, PG 18:3/16:0, PG 18:2/16:0) to accumulate preferentially in BS cells, and for FA16:1-containing PGs (i.e., PG 16:0/16:1 and PG 18:3/16:1) to accumulate preferentially in M cells, although PG 18:2/16:1 is present almost evenly between M and BS cells. These contrasts occur regardless of genotype, except for PG 16:0/16:0, which uniquely shows genotypic

differences, evenly distributed between M and BS cells in Mo17 and Mo17xB73 but more concentrated in BS cells in B73 and B73xMo17.

Table 1. Summary of cell-specific localization of phosphatidylglycerol (PG) species

	B73	B73 × Mo17	Mo17 × B73	Mo17
PG 32:1 (16:0/16:1)	M (0.66–0.72)	M (0.23–0.70)	M (0.01–0.93)	M (0.09–0.78)
PG 32:0 (16:0/16:0)	BS (2.8–9.3)	BS (1.9–2.8)	M and BS (0.70–1.4)	M and BS (0.71–1.5)
PG 34:4 (18:3/16:1)	M (0.46–0.91)	M (0.14–0.48)	M (0.15–0.44)	M (0.22–0.98)
PG 34:3 (18:2/16:1)	M and BS (1.1–2.4)	M and BS (1.2–2.8)	M and BS (0.93–1.3)	M and BS (1.1–1.4)
PG 34:3 (18:3/16:0)	BS (2.2–3.4)	BS (2.3–4.8)	BS (1.6–3.2)	BS (2.2–5.4)
PG 34:2 (18:2/16:0)	BS (5.0–83)	BS (6.4–97)	BS (2.9–87)	BS (6.9–57)

*The numbers in parenthesis represent the range of ion signal ratios in BS versus M cells per unit pixel area obtained from three replicates of midpoint sections. Detailed analysis is presented in Table S1.

Discussion

The physical separation of different cell types can be performed for lipidomic, transcriptomic or proteomic analysis, but the sample processing involved can result in undesirable degradation or cross-contamination of the materials. MALDI-MSI provides distinct advantages for the visualization of metabolite distributions within intact tissues and cellular environments. This is especially important because the molecular distributions of metabolites or lipids may not always be the same, even among the same cell types. For example, 2,4-dihydroxy-7-methoxy-1,4-

benzoxazin-3-one glucoside (DIMBOA-Glc) and 2-hydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside (HMBOA-Glc) are known to be present in M cells of maize, but MALDI-MSI revealed that they are present only in M cells between each pair of vascular bundles (Korte, et al. 2015). This is also observed in the current dataset for all four genotypes (Figures S10, S11).

The mechanisms underlying tissue development and function are fundamental questions of cellular biology. Maize is an excellent model to study the establishment of spatially distributed metabolic functions in which metabolism is facilitated by specialized cellular differentiation and cellular arrangement. In maize leaves, this is specifically associated with the differentiation of two types of photosynthetic cells, the bundle sheath and the mesophyll, that are characteristic of the Kranz anatomy (Laetsch 1974), facilitating C4 photosynthetic photoautotrophy (Edwards *et al.* 2001, Hatch 1987, von Caemmerer and Furbank 2003). The chloroplasts within BS and M cells harbor different metabolic capabilities and express different ultrastructures: the BS chloroplasts contain Rubisco and catalyze the Calvin cycle for CO_2-fixation, whereas these capabilities are absent from the chloroplasts of M cells (Nelson and Langdale 1992). In contrast, photosystem II (PSII) and granal stacks are absent from BS chloroplasts (Woo *et al.* 1970).

Taking advantage of the MSI capabilities that we have developed over the past 8 years, here we directly tested the hypothesis that the lipidomes of M and BS cells in maize are differentially affected by cellular development, and that this developmental program is genetically controlled, as determined by the state of the maize genome. Specifically, we applied high spatial-resolution MSI (performed at either 5- or 10-μm resolution) on four developmental stages from leaves of four different maize genotypes, and compared the cellular location and relative abundances of two major anionic thylakoid lipids, SQDG and PG, within mesophyll and bundle sheath cells.

Typical of monocot leaves, the cellular differentiation that establishes the Kranz anatomy is linearly arranged from the base to the tip of the maize leaf, greatly facilitating its molecular and cellular study (Nelson and Langdale 1989). Recent studies have revealed the dynamic nature of cellular differentiation as changes in the transcriptome and proteome along the developmental gradient of the leaf (Li, et al. 2010). These analyses have uncovered a programmed change in the pattern of gene expression along the developmental length of the leaf, which parallels known morphometric alterations in cellular differentiation that marks the establishment of the Kranz anatomy (Li, et al. 2010). Specifically, in the basal section of the leaf, which is associated with the zone of cell division and elongation, expressed genes are markedly enriched for functions encoding for protein, DNA and cell wall biosynthesis. In the proximal section of the leaf that is associated with the cellular transition from sink to source tissue, the expressed genes are associated with photosynthetic machinery and secondary wall biosynthesis functions. Subsequently in the distal portion of the leaf, which is fully photoautotrophic, expressed genes associated with photosynthetic reactions, including the Calvin cycle, are enriched.

Similar transcriptomic (Tausta, et al. 2014) and proteomic (Majeran, et al. 2008) studies have specifically profiled the gene expression programs of BS and M cells along the developmental gradient of the maize leaf. These studies demonstrated the dynamic differences in gene expression between the two photosynthetic cell types, with more than twice the number of differentially expressed genes at the midpoint section of the leaf, as compared with the mature leaf tip. These molecular profiling studies identified differential expression patterns between M and BS cells, including those of photosystems I and II proteins, and thylakoid and envelope membrane proteins that function as metabolite transporters (Manandhar-Shrestha, et al. 2013). These molecular differences between M and BS cells are consistent with morphological and functional

differences between these two cell types, particularly associated with the two chloroplast populations that differentiate the two cell types. These attributes particularly distinguish the photosynthetic capabilities and thylakoid membrane ultrastructures of the two chloroplast types (Hatch 1987).

Integrating the recent data on the dynamics of the transcriptomic and proteomic profiles with the ultrastructural changes associated with the developmental gradients of the two photosynthetic cell types, led to the hypothesis that there will be lipidomic differences between these two cell types and these will differentially affect the developmental gradient of the leaf. Indeed, earlier physical fractionation-based lipid analyses of isolated chloroplasts provided indications of such lipidome differences among chloroplasts of M and BS cells (Edwards and Black 1971, Nishihara, et al. 1980, Poincelot 1973, Woo, et al. 1970). However, conclusions from these earlier studies were confounded by the use of fractionation methods that included slow digestion processes, and were susceptible to cross-contamination of the M and BS cell fractions (Edwards and Black 1971, Kanai and Edwards 1973, Majeran, et al. 2005, Nelson and Langdale 1992, Poincelot 1973). They were further confounded by the use of more complex analytical technologies, which often needed enzymatic digestion with undefined reagents (e.g. digestion of cell wall components with crude fungal-derived hydrolytic enzyme preparation) or chemical modification reactions of the tissue or the isolated lipids (e.g. the use of positional-sensitive lipases or chemical reagents to characterize lipid molecular species; Christie and Han 2012, Kuksis *et al.* 1983, Siebertz *et al.* 1979). The application of modern advances in mass spectrometry has more recently provided enormous progress in the understanding of acyl lipid profiles and lipid metabolism in plants (Gasulla *et al.* 2013, Li-Beisson *et al.* 2013, Ramadan *et al.* 2014, Riedelsheimer *et al.* 2013, Tarazona *et al.* 2015, Vu *et al.* 2014, Welti *et al.* 2002).

SQDG and PG are chloroplast lipids that are crucial for maintaining chloroplast structure and function, including the stability of the PS II complex in the thylakoid membranes (Kansy *et al.* 2014, Yu and Benning 2003), and they are essential for the growth and development of photoautotrophs (Yu and Benning 2003). PG could be distinguished from SQDG by the fact that it has at least one 16-carbon fatty acyl chain, either 16:0 or 16:1. We could not determine whether these 16-carbon acyl-chains occur at the sn-1 or sn-2 position, but earlier studies indicate that both 16:0 and 16:1 prefer the sn-2 position, with 16:1 having the priority if both acyl chains are present (i.e., PG 16:0/16:1; Nishihara, et al. 1980). One clear difference between our data and the earlier study is the presence of PG 18:2/18:3 localized to BS cells (Nishihara, et al. 1980). According to the extract analysis, we do observe PG 36 species but in much lower abundance (Figure S15). Hence, the difference is most likely associated with the difference in genotype (L. var. Honey Buntum) and/or age of tissue (45-day old plants).

The relative abundances of SQDG molecular species determined by MSI is consistent with earlier bulk-extract analysis of this class of lipid: namely, 16:0- and 18:3-containing SQDG lipids account for the majority of this lipid class (Kenrick and Bishop 1986), with 80-90% being SQDG 32:0 and SQDG 34:3. There are only minor differences in the abundance of these lipids among the four genotypes evaluated; however, development dramatically affects their abundance. Particularly, the most abundant SQDG 34:2 in the basal zone is replaced by SQDG 34:3 as the leaf develops. In the third leaf that we assessed for molecular imaging, this cellular differentiation occurs as the leaf expands through the ligule of the second leaf, and becomes directly exposed to illumination. This is marked by the differentiation of proplastids to different types of chloroplasts that are non-uniformly distributed between BS and M cells, and this differentiation is associated with an increase in the size and number of chloroplasts per cell. At this differentiation of the M

and BS chloroplasts, granal stacks begin to develop in M chloroplasts, but not in BS chloroplasts (Majeran *et al.* 2010). Therefore, the observed changes in the distribution of the different molecular species of SQDG may suggest that SQDG 34:3 is mostly associated with thylakoid membranes, whereas SQDG 34:2 is a major component of the chloroplast envelopes.

In the case of PG, PG 34:2 is the dominant molecular species in the basal zone of the leaf, accounting for 80-90% of this lipid class, with essentially no difference among the inbreds and hybrids evaluated (Figure 3). As leaf development progresses, the abundance of other PG molecular species (PG 32:0 or PG 34:3) increase and they become most dominant at the midpoint and distal end sections. This is consistent with earlier work that showed the increase of 18:3 and the decrease of 18:2 fatty acyl composition of PG across the developmental gradient of the maize leaf (Roughan 1985). Using the same logic as discussed above for SQDG molecular species, these findings may suggest that PG 34:2 is mostly associated with chloroplast envelopes, and that PG 32:0 and PG 34:3 molecular species are primarily associated with thylakoid membranes. The relative abundance of PG molecular species, however, continuously changes, at least up to the midpoint (Figure 3B), whereas SQDG shows only minimal change after the proximal zone (Figure 3A). Furthermore, the 16:1 fatty acyl composition of PG continuously increases as the leaf matures, manifested by the increase of PG 32:1 (Figure 3B) and PG 34:3 (18:2/16:1) (Figure 5) from the basal to distal zones of the leaf. This suggests the importance of PG molecular species, specifically with 16 carbon fatty acyl chains, for the functional development of chloroplasts, not merely anatomical differentiation mostly dictated by 18:2 and 18:3. In addition to the developmental effect on the relative abundances among PG molecular species, there is also a genotype dependence on the distribution of these lipids. This difference among the genotypes is

especially prominent at the proximal end of the leaf, and becomes indistinguishable as the leaf expands into the mature stages.

Of significance, there is a difference in the localization of PG molecular species between BS and M cells, which appears to be dependent on the fatty acyl chains, and it is also a trait that is partially influenced by genotype. Such a heterogeneity was previously suggested by the lipid analysis of physically separated BS and M cells (Nishihara, et al. 1980). This differential localization is not exclusive but is rather preferential, and major PGs are present in both BS and M cells. As summarized in Table 1, PGs with FA16:0 and FA16:1 prefer to localize in BS and M cells, respectively, for all genotypes. The preferential localization of FA16:1-containing PGs in M cells and their increased accumulation in later stages of leaf development may indicate their importance to the development of the PSII complex and granal stacks, which are localized in M cells and most enriched in the distal region of the leaf (Li, et al. 2010, Majeran, et al. 2010). Some PGs are much more homogeneously localized between the two cell types, specifically PG 18:2/16:1 in all four genotypes (Figure 6) and PG 16:0/16:0 in Mo17 and Mo17xB73 (Figure 4). The homogeneous distribution of these PG molecular species may suggest the importance of PG 18:2/16:1 and PG 32:*x* in both BS and M cells. Considering the importance of 16:1-containing PGs, it is reasonable to hypothesize that a corresponding desaturase would be differentially expressed in large-scale proteomics or transcriptomics data sets, especially in leaf development and/or between different cell types (Li, et al. 2010, Majeran, et al. 2008). In Arabidopsis, fatty acid desaturase 4 (FAD4) is known to be involved in the formation of $\Delta^{3\text{-trans}}$ FA16:1 at the sn-2 position of PGs (Gao *et al.* 2009). The maize genome includes two homologs of FAD4 in the NCBI BLASTP search, i.e. GRMZM2G175401 and GRMZM2G097509, and both of them are expressed in shoots, especially at the tip of stage-2 leaves of V5 and V7, respectively (Sekhon *et al.* 2011).

Furthermore, GRMZM2G175401 is shown to have a higher expression in M cells than BS cells according to the supplementary data of Li et al. (2010); however, the correlation between the content of 16:1t in PGs and FAD4 is obscure as the reduction of 16:1t has been also reported for other mutants in PG biosynthesis such as *pgpp1-1* and *pgp-1* (Babiychuk *et al.* 2003, Kobayashi *et al.* 2015, Lin *et al.* 2016).

In summary, the work described herein demonstrates that high-resolution MALDI-MSI analysis can be directly applied to multicellular plant tissues to uncover cell-specific metabolic biology that has not been possible using traditional metabolomics methodology. This capability of *in situ* imaging is specifically valuable for metabolites and lipids because of their highly dynamic nature compared with transcriptome or proteome. Combined with other single cell '-omics' technologies (Dai and Chen 2012, Shalek *et al.* 2013, Wang and Bodovitz 2010), high-resolution MSI will allow for enhanced systems biological understanding at the single-cell level in the foreseeable future.

Experimental Procedures

Plant growth and sample preparation

The overall experimental workflow is illustrated in Figure S1. Maize kernels (*Zea mays* L. inbreds B73 and Mo17, and reciprocal hybrids B73xMo17 and Mo17xB73) were planted in soil, and grown in a climate-controlled glasshouse at 30% humidity under a diurnal cycle of 16 h of light and 8 h of dark at 27°C and 24°C, respectively. Seedlings were harvested at approximately 13:00 h local time, 11-13 days after planting, when the length of the third leaf was 14-18 cm, as measured from the second ligule. Sections from four positions along the developmental gradient of leaf 3 (Figure S2) were collected from three individual plants of each genotype, at similar positions as described by Li *et al* (2010). The etiolated basal zone leaf section was collected as a

cross-section through the sheathed basal zone of leaf 3 within the whorl, 1.5 cm below leaf ligule 2. The second section of leaf 3, labeled as the proximal section, was collected 1.5 cm above the position of the ligule 2 and at the point of emergence from the whorl, where the leaf begins to unfold and becomes fully exposed to direct illumination. The third section of leaf 3, labeled as the midpoint, was collected halfway between the proximal section and the leaf tip. The fourth section of leaf 3, labeled as the distal section, was collected ~2 cm from the tip of the leaf and approximately 12-16 cm from the base.

Tissue samples were cryosectioned and prepared for MALDI-MSI as described previously (Korte and Lee 2014). Briefly, fresh maize leaf sections were placed in a cryomold, then submerged in gelatin (10% w/v solution) and immediately frozen with liquid nitrogen. The molds were then transferred to a cryostat (CM1850, Leica Microsystems; Buffalo Grove, IL, USA) that was pre-chilled to -20°C, and the samples were allowed to thermally equilibrate for 30 minutes. Leaf tissue was cryo-sectioned at 10 μm thickness and was then collected with Cryo-Jane tape (Leica Biosystems, http://www.leicabiosystems.com), and attached to pre-chilled glass slides. The prepared slides were placed onto a chilled aluminum block and were vacuum dried while gradually warming to room temperature. After acquiring optical microscope images (Axio Zoom.V16; Carl Zeiss, http://www.zeiss.com), the dried tissues were subjected to matrix deposition by sublimating 1,5-diaminonaphthalene (97%; Sigma-Aldrich, http://www.sigmaaldrich.com) (Hankin *et al.* 2007) at 140°C for 4 minutes at a pressure of ~50 mtorr.

For the fluorescence image shown in Figure S12, a leaf tissue was cryosectioned at 10-μm thickness and collected on a charged glass slide (Fisher Scientific, http://www.fishersci.com) by thaw mounting. The fluorescence image was obtained with an Olympus BX-60 microscope (http://www.olympus-global.com) with a Jenoptik C-5 camera (http://www.jenoptik.com). A

consecutive cross-section was collected with Cryo-Jane tape to compare with the MS image of chlorophyll *a*.

Mass spectrometry analysis

Mass spectrometry imaging data were collected using a MALDI-linear ion trap (LIT)-Orbitrap mass spectrometer (MALDI-LTQ-Orbitrap Discovery; Thermo Scientific, http://www.thermofisher.com). The instrument was modified to use an external 355-nm frequency tripled Nd: YAG laser (UVFQ; Elforlight Ltd., http://www.elforlight.com). The laser energy used was 83-84% (~1 μJ per pulse) at a 60 Hz repetition rate. The laser optics are similar to those previously described by Korte *et al.* (2015), but the focus lens was further reduced to 60 or 75 mm and a commercial 10X or 5X beam expander (Thorlabs, http://www.thorlabs.com) was used to reduce the laser spot size. The laser spot size was 9-11 μm for 10 μm-resolution imaging with a 5X beam expander and 6-7 μm for 5 μm-resolution imaging with a 10X beam expander; thus, 5-μm resolution imaging was acquired using an oversampling method. TunePlus and Xcalibur (ThermoFisher Scientific) were used to define imaging parameters and to acquire data, respectively. Mass spectra were acquired with 10 laser shots per spectrum in negative mode using an Orbitrap mass analyzer (resolution of 30 000 at *m/z* 400) for an *m/z* scan range of 100-1000.

All lipids were detected as deprotonated, $[M-H]^-$. MS images were generated using ImageQuest (ThermoFisher Scientific) with a mass window of ±0.003 Da, and with normalization to the total ion count (TIC) for Figures 1, 2, and S4-S7, and without normalization for the rest of the figures. Overlays of MS images with the optical microscope images were made using the software module of MATLAB, Image Processing Toolbox 9.3 (MATLAB 2015). MSiReader v.0.09 (Robichaud *et al.* 2013) was used to obtain the ion signal ratios of PG molecular species between BS and M cells normalized to each cell area. First, the optical images and MS images

were overlaid using the MSiImage tool. Then, the region of interest (ROI) for the bundle sheath and mesophyll cells was selected using the polygon drawing tool and the spectra intensity data for each ROI was exported for all PG species. The total intensity of each PG species at each ROI was then normalized to the area of the ROI for each cell type.

MS/MS imaging was performed using the ion trap analyzer for selected ions using the same conditions as described for MS imaging. An isolation width of 2.0 Da and normalized collision energy of 35 were used. MS/MS images were generated using MSiReader with a mass window of ±0.2 Da and no normalization to the TIC. Quantitative comparison was made by averaging metabolite intensities over the entire imaging area of the maize leaf that contained both mesophyll and bundle sheath cells. Three or more biological replicates of each genotype were analyzed, and the average is reported with standard deviation. Student's t-tests were performed to calculate pairwise P-values.

Lipid Extraction and Direct-Infusion ESI-MS

Total lipids were extracted from maize leaf (inbred B73) as described by Shiva *et al.* (2013). Lipid samples were then purified by two-phase partitioning with chloroform and 1 M KCl, dried under N_2 and resuspended in chloroform, and stored at -20°C until MS analysis. Three biological replicates were analyzed, and the average is reported with standard deviation. Total lipid extracts were analyzed by direct-infusion ESI-MS on a Fourier Transform Ion Cyclotron Resonance mass spectrometer (7T SolariX FT-ICR MS; Bruker Daltonics, http://www.bruker.com).

Acknowledgements

This work was supported by the US Department of Energy (DOE), Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences. MDY-N and BJN

acknowledge the support of the National Science Foundation under Award No. EEC-0813570 and Award No. IOS-1354799, which co-sponsored the development of the genetic stocks imaged in this study. The Ames Laboratory is operated by Iowa State University under DOE Contract DE-AC02-07CH11358.

References

Angel, P.M. and Caprioli, R.M. (2013) Matrix-assisted laser desorption ionization imaging mass spectrometry: in situ molecular mapping. *Biochemistry*, **52**, 3818-3828.

Babiychuk, E., Müller, F., Eubel, H., Braun, H.-P., Frentzen, M. and Kushnir, S. (2003) Arabidopsis phosphatidylglycerophosphate synthase 1 is essential for chloroplast differentiation, but is dispensable for mitochondrial function. *The Plant Journal*, **33**, 899-909.

Ben Hamed, K., Ben Youssef, N., Ranieri, A., Zarrouk, M. and Abdelly, C. (2005) Changes in content and fatty acid profiles of total lipids and sulfolipids in the halophyte Crithmum maritimum under salt stress. *J. Plant Physiol.*, **162**, 599-602.

Brown, W.V. (1975) Variations in Anatomy, Associations, and Origins of Kranz Tissue. *Am. J. Bot.*, **62**, 395-402.

Christie, W.W. and Han, X. (2012) Chapter 12 - Positional distributions of fatty acids in glycerolipids. In *Lipid Analysis (Fourth edition)*: Woodhead Publishing, pp. 261-273.

Dai, S.J. and Chen, S.X. (2012) Single-cell-type Proteomics: Toward a Holistic Understanding of Plant Function. *Mol. Cell. Proteomics*, **11**, 1622-1630.

Demé, B., Cataye, C., Block, M.A., Maréchal, E. and Jouhet, J. (2014) Contribution of galactoglycerolipids to the 3-dimensional architecture of thylakoids. *The FASEB Journal*.

Duchêne, S. and Siegenthaler, P.-A. (2000) Do glycerolipids display lateral heterogeneity in the thylakoid membrane? *Lipids*, **35**, 739-744.

Edwards, G.E. and Black, C.C. (1971) Isolation of Mesophyll Cells and Bundle Sheath Cells from Digitaria sanguinalis (L.) Scop. Leaves and a Scanning Microscopy Study of the Internal Leaf Cell Morphology. *Plant Physiol.*, **47**, 149-156.

Edwards, G.E., Franceschi, V.R., Ku, M.S.B., Voznesenskaya, E.V., Pyankov, V.I. and Andreo, C.S. (2001) Compartmentation of photosynthesis in cells and tissues of C4 plants. *J. Exp. Bot.*, **52**, 577-590.

Evert, R.F., Russin, W.A. and Bosabalidis, A.M. (1996) Anatomical and Ultrastructural Changes Associated with Sink-to-Source Transition in Developing Maize Leaves. *Int. J. Plant Sci.*, **157**, 247-261.

Gao, J.P., Ajjawi, I., Manoli, A., Sawin, A., Xu, C.C., Froehlich, J.E., Last, R.L. and Benning, C. (2009) FATTY ACID DESATURASE4 of Arabidopsis encodes a protein distinct from characterized fatty acid desaturases. *Plant J.*, **60**, 832-839.

Gasulla, F., vom Dorp, K., Dombrink, I., Zaehringer, U., Gisch, N., Doermann, P. and Bartels, D. (2013) The role of lipid metabolism in the acquisition of desiccation tolerance in Craterostigma plantagineum: a comparative approach. *Plant J.*, **75**, 726-741.

Guskov, A., Kern, J., Gabdulkhakov, A., Broser, M., Zouni, A. and Saenger, W. (2009) Cyanobacterial photosystem II at 2.9-A resolution and the role of quinones, lipids, channels and chloride. *Nat. Struct. Mol. Biol.*, **16**, 334-342.

Haberlandt, G. (1882) Vergleichende anatomie des assimilatorischen gewebesystems der pflanzen. *Jb. wiss. Bot.(13)*, 74-188.

Han, X. and Gross, R.W. (1994) Electrospray ionization mass spectroscopic analysis of human erythrocyte plasma membrane phospholipids. *Proceedings of the National Academy of Sciences*, **91**, 10635-10639.

Han, X., Yang, K. and Gross, R.W. (2012) Multi-dimensional mass spectrometry-based shotgun lipidomics and novel strategies for lipidomic analyses. *Mass Spectrom. Rev.*, **31**, 134-178.

Hankin, J., Barkley, R. and Murphy, R. (2007) Sublimation as a method of matrix application for mass spectrometric imaging. *J. Am. Soc. Mass Spectrom.*, **18**, 1646-1652.

Hatch, M.D. (1987) C4 photosynthesis: a unique elend of modified biochemistry, anatomy and ultrastructure. *Biochimica et Biophysica Acta (BBA) - Reviews on Bioenergetics*, **895**, 81-106.

Horn, P.J., Korte, A.R., Neogi, P.B., Love, E., Fuchs, J., Strupat, K., Borisjuk, L., Shulaev, V., Lee, Y.-J. and Chapman, K.D. (2012) Spatial Mapping of Lipids at Cellular Resolution in Embryos of Cotton. *Plant Cell*, **24**, 622-636.

Hsu, F.-F., Turk, J., Williams, T.D. and Welti, R. (2007) Electrospray Ionization Multiple Stage Quadrupole Ion-Trap and Tandem Quadrupole Mass Spectrometric Studies on Phosphatidylglycerol from Arabidopsis Leaves. *J. Am. Soc. Mass Spectrom.*, **18**, 783-790.

Jordan, P., Fromme, P., Witt, H.T., Klukas, O., Saenger, W. and Krausz, N. (2001) Three-dimensional structure of cyanobacterial photosystem I at 2.5[thinsp][angst] resolution. *Nature*, **411**, 909-917.

Jungmann, J.H. and Heeren, R.M. (2012) Emerging technologies in mass spectrometry imaging. *J. Proteomics*, **75**, 5077-5092.

Kanai, R. and Edwards, G.E. (1973) Separation of Mesophyll Protoplasts and Bundle Sheath Cells from Maize Leaves for Photosynthetic Studies. *Plant Physiol.*, **51**, 1133-1137.

Kaniuga, Z., Sączyńska, V., Miśkiewicz, E. and Garstka, M. (1999) The Fatty Acid Composition of Phosphatidylglycerol and Sulfoquinovosyldiacylglycerol of Zea mays Genotypes Differing in Chilling Susceptibility. *J. Plant Physiol.*, **154**, 256-263.

Kansy, M., Wilhelm, C. and Goss, R. (2014) Influence of thylakoid membrane lipids on the structure and function of the plant photosystem II core complex. *Planta*, **240**, 781-796.

Kenrick, J.R. and Bishop, D.G. (1986) The fatty acid composition of phosphatidylglycerol and sulfoquinovosyldiacylglycerol of higher plants in relation to chilling sensitivity. *Plant physiology*, **81**, 946-949.

Kirchanski, S.J. (1975) The Ultrastructural Development of the Dimorphic Plastids of Zea mys L. *Am. J. Bot.*, **62**, 695-705.

Kirchhoff, H., Sharpe, R.M., Herbstova, M., Yarbrough, R. and Edwards, G.E. (2013) Differential Mobility of Pigment-Protein Complexes in Granal and Agranal Thylakoid Membranes of C3 and C4 Plants. *Plant Physiol.*, **161**, 497-507.

Kobayashi, K., Fujii, S., Sato, M., Toyooka, K. and Wada, H. (2015) Specific role of phosphatidylglycerol and functional overlaps with other thylakoid lipids in Arabidopsis chloroplast biogenesis. *Plant Cell Rep.*, **34**, 631-642.

Kobayashi, K., Nakamura, Y. and Ohta, H. (2009) Type A and type B monogalactosyldiacylglycerol synthases are spatially and functionally separated in the plastids of higher plants. *Plant Physiol. Biochem.*, **47**, 518-525.

Korte, A., Yandeau-Nelson, M., Nikolau, B. and Lee, Y. (2015) Subcellular-level resolution MALDI-MS imaging of maize leaf metabolites by MALDI-linear ion trap-Orbitrap mass spectrometer. *Anal. Bioanal. Chem.*, **407**, 2301-2309.

Korte, A.R. and Lee, Y.J. (2014) MALDI- MS analysis and imaging of small molecule metabolites with 1, 5- diaminonaphthalene (DAN). *J. Mass Spectrom.*, **49**, 737-741.

Kuksis, A., Marai, L. and Myher, J.J. (1983) Strategy of glycerolipid separation and quantitation by complementary analytical techniques. *Journal of Chromatography B: Biomedical Sciences and Applications*, **273**, 43-66.

Laetsch, W.M. (1974) The C4 syndrome: a structural analysis. *Annual review of plant physiology*, **25**, 27.

Lee, Y.J., Perdian, D.C., Song, Z., Yeung, E.S. and Nikolau, B.J. (2012) Use of mass spectrometry for imaging metabolites in plants. *The Plant Journal*, **70**, 81-95.

Leech, R.M., Rumsby, M.G. and Thomson, W.W. (1973) Plastid Differentiation, Acyl Lipid, and Fatty Acid Changes in Developing Green Maize Leaves. *Plant Physiol.*, **52**, 240-245.

Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S.L., Kebrom, T.H., Provart, N., Patel, R., Myers, C.R., Reidel, E.J., Turgeon, R., Liu, P., Sun, Q., Nelson, T. and Brutnell, T.P. (2010) The developmental dynamics of the maize leaf transcriptome. *Nat. Genet.*, **42**, 1060-1067.

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M.X., Arondel, V., Bates, P.D., Baud, S., Bird, D., Debono, A., Durrett, T.P., Franke, R.B., Graham, I.A., Katayama, K., Kelly, A.A., Larson, T., Markham, J.E., Miquel, M., Molina, I., Nishida, I., Rowland, O., Samuels, L., Schmid, K.M., Wada, H., Welti, R., Xu, C., Zallot, R. and Ohlrogge, J. (2013) Acyl-lipid metabolism. *The Arabidopsis book / American Society of Plant Biologists*, **11**, e0161-e0161.

Lin, Y.-C., Kobayashi, K., Hung, C.-H., Wada, H. and Nakamura, Y. (2016) Arabidopsis PHOSPHATIDYLGLYCEROPHOSPHATE PHOSPHATASE1 (PGPP1) Involved in Phosphatidylglycerol Biosynthesis and Photosynthetic Function. *The Plant Journal*, n/a-n/a.

Majeran, W., Cai, Y., Sun, Q. and van Wijk, K.J. (2005) Functional Differentiation of Bundle Sheath and Mesophyll Maize Chloroplasts Determined by Comparative Proteomics. *Plant Cell*, **17**, 3111-3140.

Majeran, W., Friso, G., Ponnala, L., Connolly, B., Huang, M., Reidel, E., Zhang, C., Asakura, Y., Bhuiyan, N.H., Sun, Q., Turgeon, R. and van Wijk, K.J. (2010) Structural and Metabolic Transitions of C-4 Leaf Development and Differentiation Defined by Microscopy and Quantitative Proteomics in Maize. *Plant Cell*, **22**, 3509-3542.

Majeran, W. and van Wijk, K.J. (2009) Cell-type-specific differentiation of chloroplasts in C4 plants. *Trends Plant Sci.*, **14**, 100-109.

Majeran, W., Zybailov, B., Ytterberg, A.J., Dunsmore, J., Sun, Q. and van Wijk, K.J. (2008) Consequences of C(4) Differentiation for Chloroplast Membrane Proteomes in Maize Mesophyll and Bundle Sheath Cells. *Molecular & Cellular Proteomics : MCP*, **7**, 1609-1638.

Manandhar-Shrestha, K., Tamot, B., Pratt, E.P.S., Saitie, S.M., Braeutigam, A., Weber, A.P.M. and Hoffmann-Benning, S. (2013) Comparative proteomics of chloroplasts envelopes from bundle sheath and mesophyll chloroplasts reveals novel membrane proteins with a possible role in C4-related metabolite fluxes and development. *Frontiers in Plant Science*, **4**.

MATLAB (2015) MATLAB version 8.6.0 (R2015b). Natick, Massachusetts: The MathWorks Inc.

Murata, N., Yamada, M., Nishida, I., Okuyama, H., Sekiya, J., Hajime, W., Xu, Y.N., Wang, Z.N., Jiang, G.Z., Li, L.B. and Kuang, T.Y. (2003) Determination of the Effects of Temperatures on Phosphatidylglycerol Biosynthesis in Thylakoid Membranes by Analysis of Molecular Species. In *Advanced Research on Plant Lipids*: Springer Netherlands, pp. 171-174.

Nelson, T. and Langdale, J.A. (1992) Developmental Genetics of C4 Photosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **43**, 25-47.

Nelson, T. and Langdale, J.A.C.F.p.d.J. (1989) Patterns of Leaf Development in C4 Plants. *Plant Cell*, **1**, 3-13.

Nishihara, M., Yokota, K. and Kito, M. (1980) Lipid molecular species composition of thylakoid membranes. *Biochim. Biophys. Acta*, **617**, 12-19.

Pál, M., Leskó, K., Janda, T., Páldi, E. and Szalai, G. (2007) Cadmium-induced changes in the membrane lipid composition of maize plants. *Cereal Research Communications*, **35**, 1631-1642.

Pick, T.R., Bräutigam, A., Schlüter, U., Denton, A.K., Colmsee, C., Scholz, U., Fahnenstich, H., Pieruschka, R., Rascher, U., Sonnewald, U. and Weber, A.P.M. (2011) Systems Analysis of a Maize Leaf Developmental Gradient Redefines the Current C(4) Model and Provides Candidates for Regulation. *Plant Cell*, **23**, 4208-4220.

Poincelot, R.P. (1973) Differences in Lipid Composition between Undifferentiated and Mature Maize Chloroplasts. *Plant Physiol.*, **51**, 802-804.

Quinn, P. and Williams, W. (1983) The structural role of lipids in photosynthetic membranes. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, **737**, 223-266.

Ramadan, A., Sabir, J.S.M., Alakilli, S.Y.M., Shokry, A.M., Gadalla, N.O., Edris, S., Al-Kordy, M.A., Al-Zahrani, H.S., El-Domyati, F.M., Bahieldin, A., Baker, N.R., Willmitzer, L. and Irgang, S. (2014) Metabolomic Response of Calotropis procera Growing in the Desert to Changes in Water Availability. *PLoS One*, **9**.

Riedelsheimer, C., Brotman, Y., Méret, M., Melchinger, A.E. and Willmitzer, L. (2013) The maize leaf lipidome shows multilevel genetic control and high predictive value for agronomic traits. *Sci. Rep.*, **3**, 2479.

Robichaud, G., Garrard, K.P., Barry, J.A. and Muddiman, D.C. (2013) MSiReader: an open-source interface to view and analyze high resolving power MS imaging files on Matlab platform. *J. Am. Soc. Mass Spectrom.*, **24**, 718-721.

Roughan, P.G. (1985) Phosphatidylglycerol and chilling sensitivity in plants. *Plant Physiol.*, **77**, 740-746.

Sakurai, I., Shen, J.-R., Leng, J., Ohashi, S., Kobayashi, M. and Wada, H. (2006) Lipids in Oxygen-Evolving Photosystem II Complexes of Cyanobacteria and Higher Plants. *J. Biochem.*, **140**, 201-209.

Sato, N. (2004) Roles of the acidic lipids sulfoquinovosyl diacylglycerol and phosphatidylglycerol in photosynthesis: their specificity and evolution. *J. Plant Res.*, **117**, 495-505.

Sekhon, R.S., Lin, H., Childs, K.L., Hansey, C.N., Buell, C.R., de Leon, N. and Kaeppler, S.M. (2011) Genome-wide atlas of transcription during maize development. *The Plant Journal*, **66**, 553-563.

Shalek, A.K., Satija, R., Adiconis, X., Gertner, R.S., Gaublomme, J.T., Raychowdhury, R., Schwartz, S., Yosef, N., Malboeuf, C., Lu, D.N., Trombetta, J.J., Gennert, D., Gnirke, A., Goren, A., Hacohen, N., Levin, J.Z., Park, H. and Regev, A. (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature*, **498**, 236-240.

Sharpe, R.M., Mahajan, A., Takacs, E.M., Stern, D.B. and Cahoon, A.B. (2011) Developmental and cell type characterization of bundle sheath and mesophyll chloroplast transcript abundance in maize. *Curr. Genet.*, **57**, 89-102.

Shiva, S., Vu, H.S., Roth, M.R., Zhou, Z., Marepally, S.R., Nune, D.S., Lushington, G.H., Visvanathan, M. and Welti, R. (2013) Lipidomic Analysis of Plant Membrane Lipids by Direct Infusion Tandem Mass Spectrometry. In *Plant Lipid Signaling Protocols*. Totowa, NJ: Humana Press, pp. 79-91.

Siebertz, H.P., Heinz, E., Linscheid, M., Joyard, J. and Douce, R. (1979) Characterization of Lipids from Chloroplast Envelopes. *Eur. J. Biochem.*, **101**, 429-438.

Sturtevant, D., Lee, Y.-J. and Chapman, K.D. (2016) Matrix assisted laser desorption/ionization-mass spectrometry imaging (MALDI-MSI) for direct visualization of plant metabolites in situ. *Curr. Opin. Biotechnol.*, **37**, 53-60.

Svatos, A. (2010) Mass spectrometric imaging of small molecules. *Trends Biotechnol.*, **28**, 425-434.

Tarazona, P., Feussner, K. and Feussner, I. (2015) An enhanced plant lipidomics method based on multiplexed liquid chromatography-mass spectrometry reveals additional insights into cold- and drought-induced membrane remodeling. *Plant J.*, **84**, 621-633.

Tausta, S.L., Li, P., Si, Y., Gandotra, N., Liu, P., Sun, Q., Brutnell, T.P. and Nelson, T. (2014) Developmental dynamics of Kranz cell transcriptional specificity in maize leaf reveals early onset of C(4)-related processes. *J. Exp. Bot.*, **65**, 3543-3555.

Umena, Y., Kawakami, K., Shen, J.-R. and Kamiya, N. (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9[thinsp]A. *Nature*, **473**, 55-60.

van Hove, E.R.A., Smith, D.F. and Heeren, R.M. (2010) A concise review of mass spectrometry imaging. *J. Chromatogr.*, **1217**, 3946-3954.

von Caemmerer, S. and Furbank, R.T. (2003) The C4 pathway: an efficient CO2 pump. *Photosynthesis Res.*, **77**, 191-207.

Vu, H.S., Shiva, S., Roth, M.R., Tamura, P., Zheng, L., Li, M., Sarowar, S., Honey, S., McEllhiney, D., Hinkes, P., Seib, L., Williams, T.D., Gadbury, G., Wang, X., Shah, J. and Welti, R. (2014) Lipid changes after leaf wounding in Arabidopsis thaliana: expanded lipidomic data form the basis for lipid co-occurrence analysis. *The Plant Journal*, **80**, 728-743.

Wang, D.J. and Bodovitz, S. (2010) Single cell analysis: the new frontier in 'omics'. *Trends Biotechnol.*, **28**, 281-290.

Wang, L., Czedik-Eysenberg, A., Mertz, R.A., Si, Y., Tohge, T., Nunes-Nesi, A., Arrivault, S., Dedow, L.K., Bryant, D.W., Zhou, W., Xu, J., Weissmann, S., Studer, A., Li, P., Zhang, C., LaRue, T., Shao, Y., Ding, Z., Sun, Q., Patel, R.V., Turgeon, R., Zhu, X., Provart, N.J., Mockler, T.C., Fernie, A.R., Stitt, M., Liu, P. and Brutnell, T.P. (2014) Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. *Nat Biotech*, **32**, 1158-1165.

Welti, R., Li, W.Q., Li, M.Y., Sang, Y.M., Biesiada, H., Zhou, H.E., Rajashekar, C.B., Williams, T.D. and Wang, X.M. (2002) Profiling membrane lipids in plant stress responses - Role of phospholipase D alpha in freezing-induced lipid changes in Arabidopsis. *J. Biol. Chem.*, **277**, 31994-32002.

Woo, K.C., Anderson, J.M., Boardman, N.K., Downton, W.J.S., Osmond, C.B. and Thorne, S.W. (1970) Deficient Photosystem II in Agranal Bundle Sheath Chloroplasts of C(4) Plants. *Proc. Natl. Acad. Sci. U. S. A.*, **67**, 18-25.

Yu, B. and Benning, C. (2003) Anionic lipids are required for chloroplast structure and function in Arabidopsis. *The Plant Journal*, **36**, 762-770.

Zaima, N., Hayasaka, T., Goto-Inoue, N. and Setou, M. (2010) Matrix-assisted laser desorption/ionization imaging mass spectrometry. *International journal of molecular sciences*, **11**, 5040-5055.

Zavalin, A., Yang, J., Hayden, K., Vestal, M. and Caprioli, R.M. (2015) Tissue protein imaging at 1 μm laser spot diameter for high spatial resolution and high imaging speed using transmission geometry MALDI TOF MS. *Anal. Bioanal. Chem.*, **407**, 2337-2342.

Supporting Information

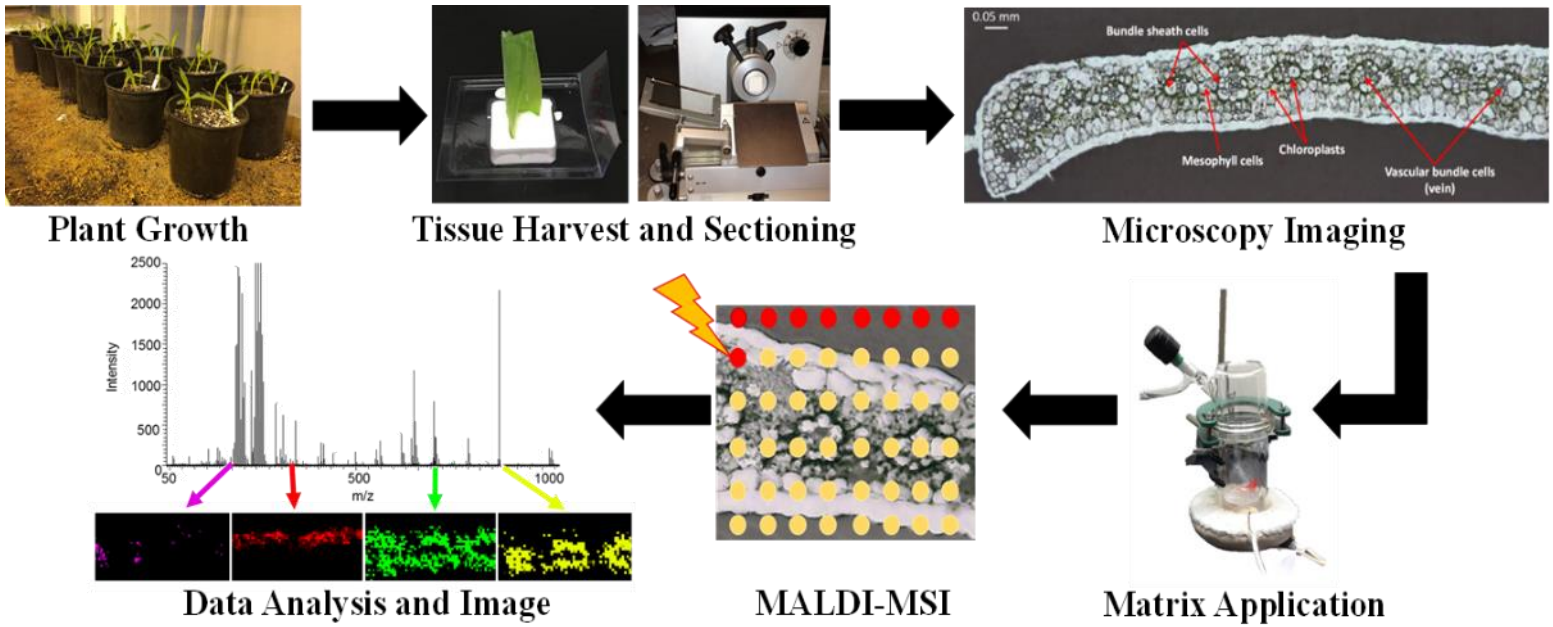


Figure S1. Overall workflow for MALDI-MSI of maize leaves

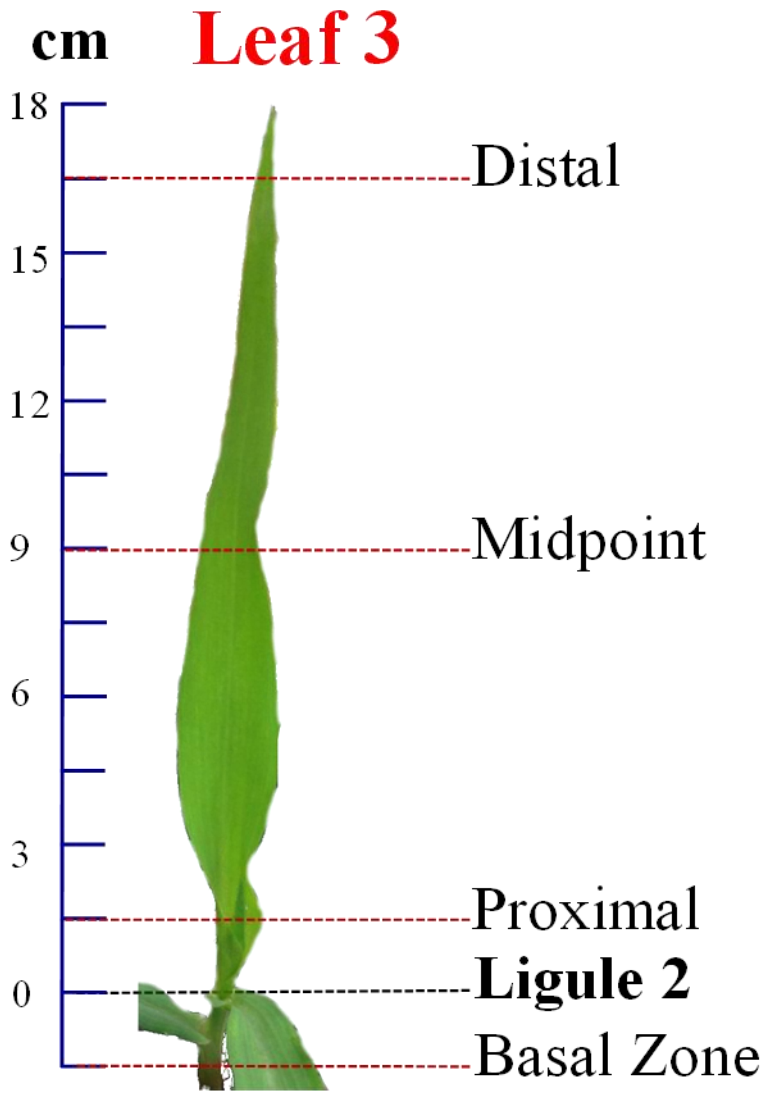


Figure S2. Illustration of four positions along the developmental gradient of maize leaf 3 used in this study. The distal section corresponds to 70~90% maturation when compared to the unified development model (UDM) based on the length, while the midpoint section collected at 6-8 cm corresponds to 36-47% maturation. The proximal section, in contrast, is expected to correspond to only ~15% maturation.

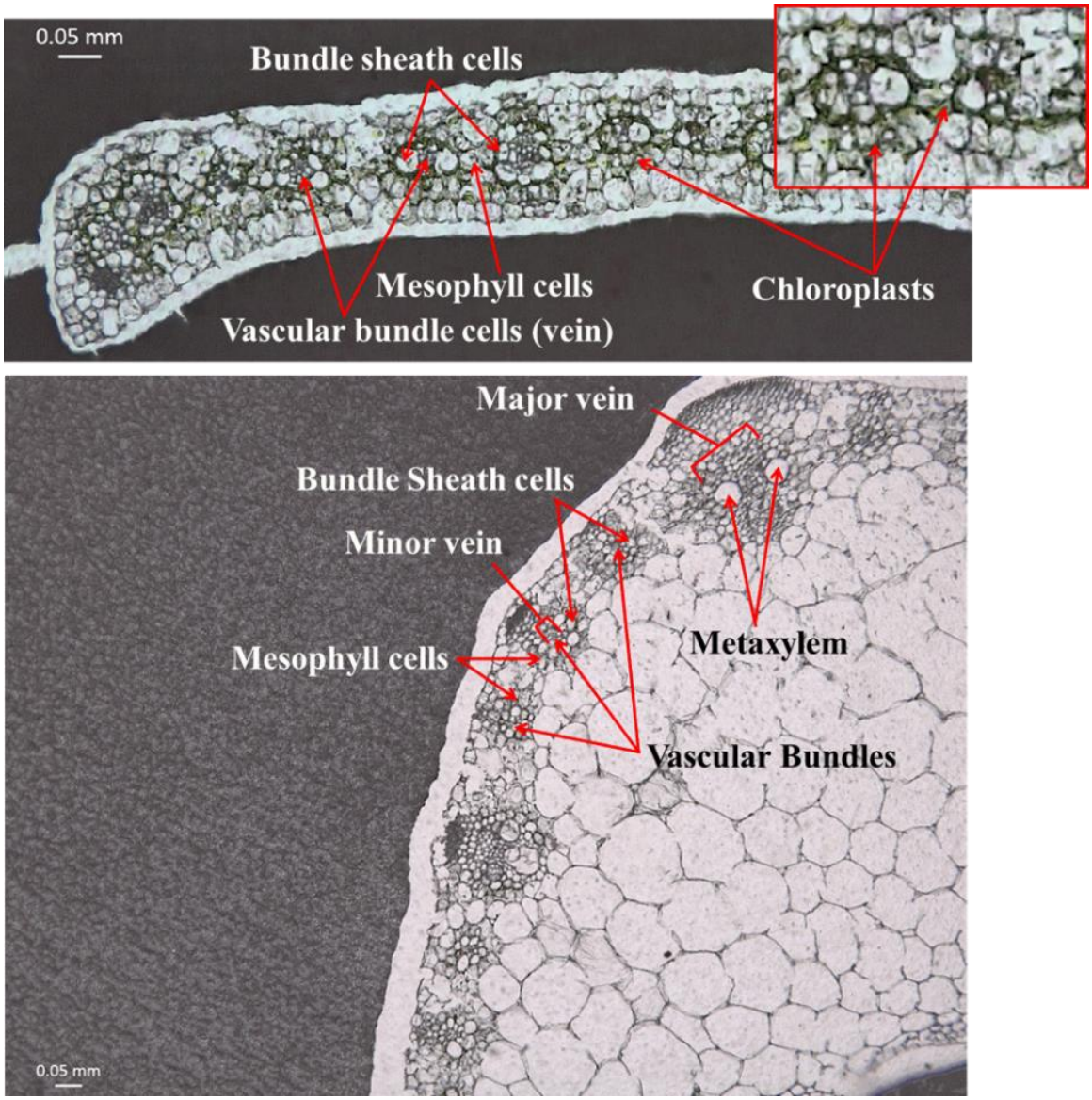


Figure S3. Transverse sections of a B73 maize leaf (distal and proximal sections, respectively) showing the characteristic C4 leaf Kranz anatomy where concentric wreaths of mesophyll and bundle sheath cells surround closely spaced veins, the vascular bundles.

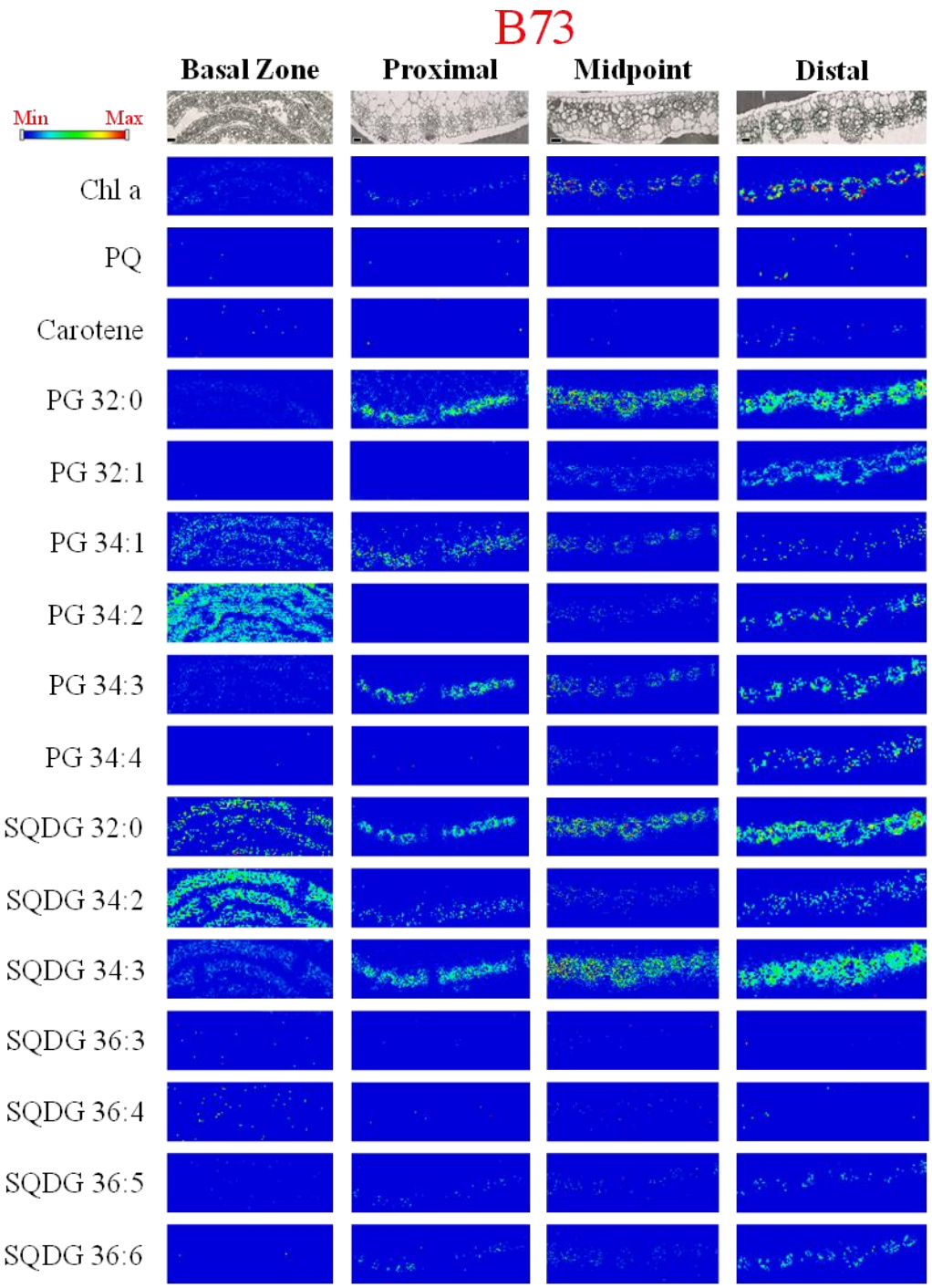


Figure S4. MS images of PGs, SQDGs, and three photosynthesis-related metabolites at the basal zone, proximal, midpoint, and distal sections of B73 inbred maize leaf. Each maximum value for color scheme was adjusted for the best comparison across the sections and between the genotypes. Chl a: chlorophyll *a* [sum of fragments at m/z 591.261 (pheophorbide *a*), 613.232 (chlorophyllide *a*), and 870.566 (pheophytin *a*)]. PQ: plastoquinone/plastoquinol. Scale bar represents 50 µm.

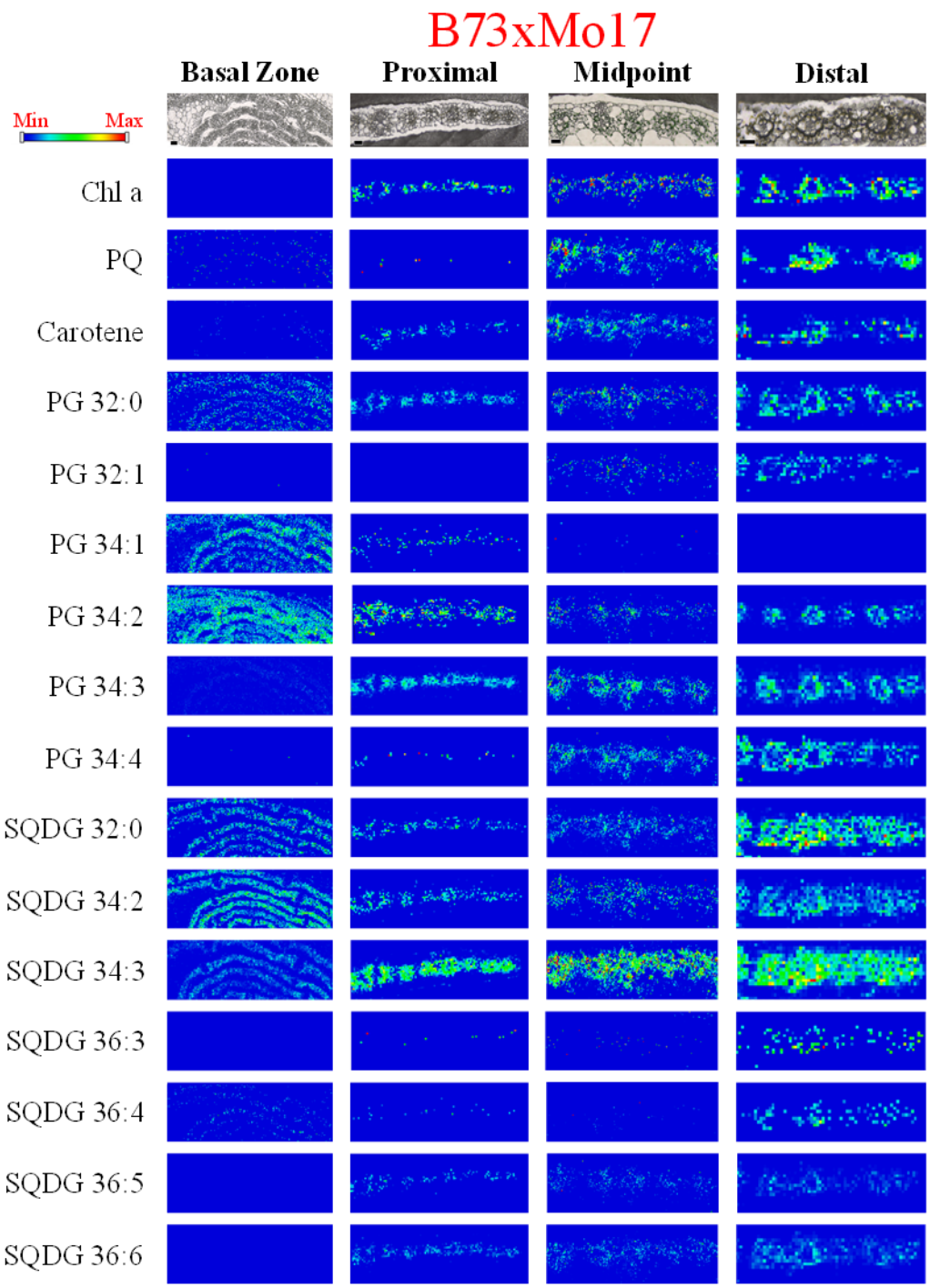


Figure S5 MS images of PGs, SQDGs, and three photosynthesis related metabolites at the basal zone, proximal, midpoint, and distal sections of B73 x Mo17 hybrid maize leaf. Each maximum value for color scheme was adjusted for the best comparison across the sections and between the genotypes. Chl a: chlorophyll *a* [sum of fragments at m/z 591.261 (pheophorbide *a*), 613.232 (chlorophyllide *a*), and 870.566 (pheophytin *a*)]. PQ: plastoquinone/plastoquinol. Scale bar represents 50 µm.

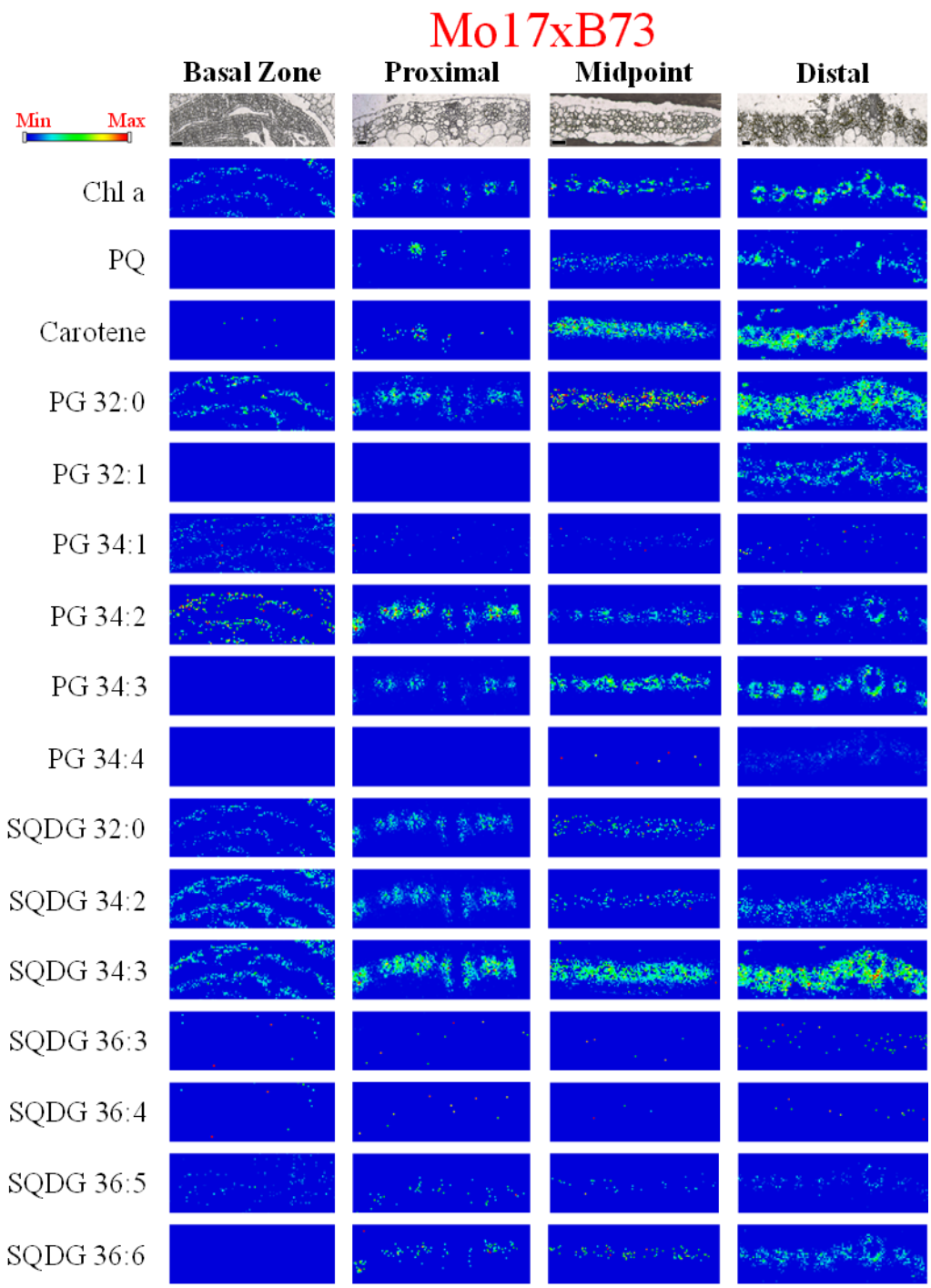


Figure S6. MS images of PGs, SQDGs, and three photosynthesis related metabolites at the basal zone, proximal, midpoint, and distal sections of Mo17 x B73 hybrid maize leaf. Each maximum value for color scheme was adjusted for the best comparison across the sections and between the genotypes. Chl a: chlorophyll *a* [sum of fragments at m/z 591.261 (pheophorbide *a*), 613.232 (chlorophyllide *a*), and 870.566 (pheophytin *a*)]. PQ: plastoquinone/plastoquinol. Scale bar represents 50 µm.

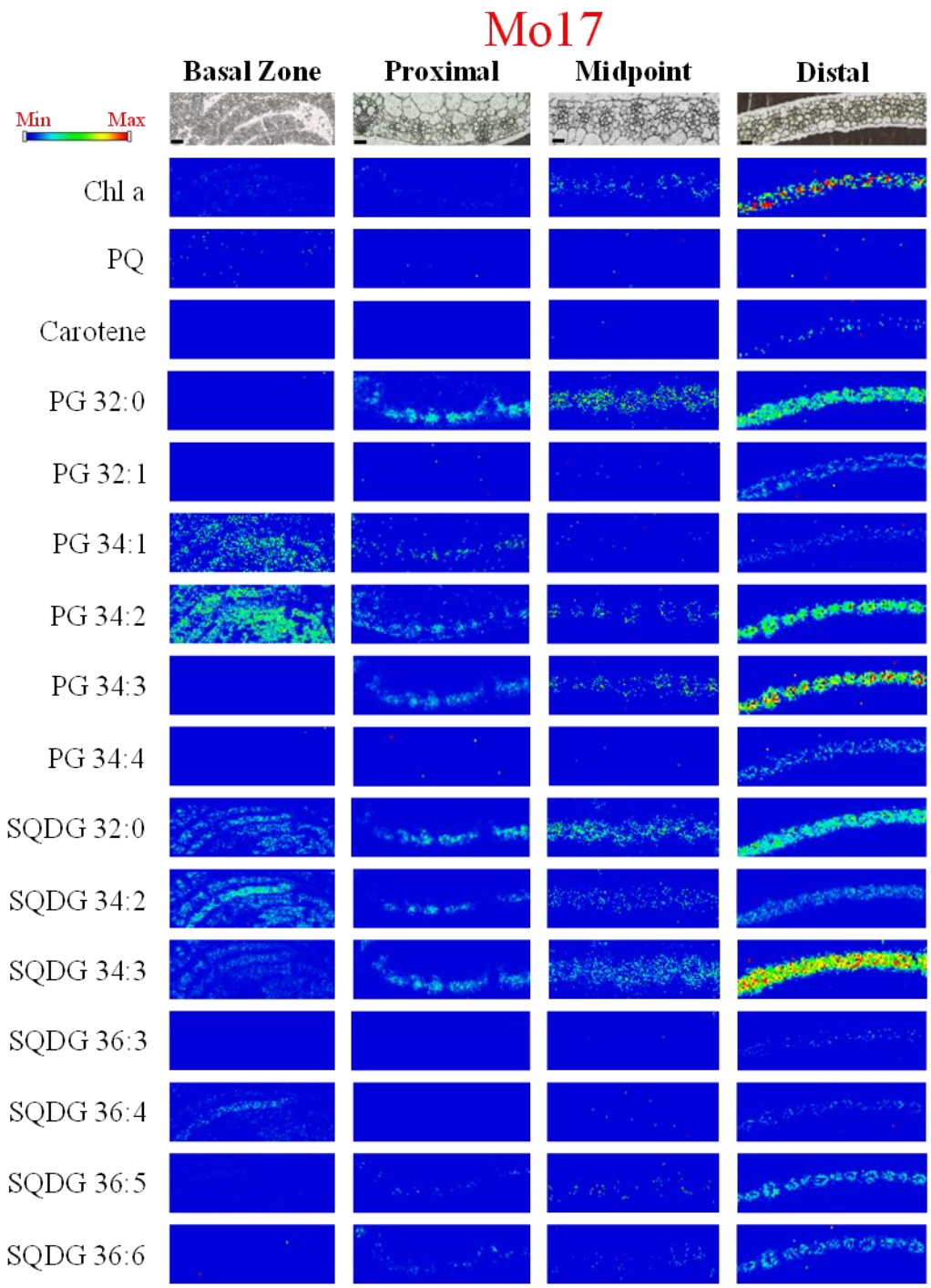


Figure S7. MS images of PGs, SQDGs, and three photosynthesis related metabolites at the basal zone, proximal, midpoint, and distal sections of Mo17 inbred maize leaf. Each maximum value for color scheme was adjusted for the best comparison across the sections and between the genotypes. Chl a: chlorophyll *a* [sum of fragments at m/z 591.261 (pheophorbide *a*), 613.232 (chlorophyllide *a*), and 870.566 (pheophytin *a*)]. PQ: plastoquinone/plastoquinol. Scale bar represents 50 µm

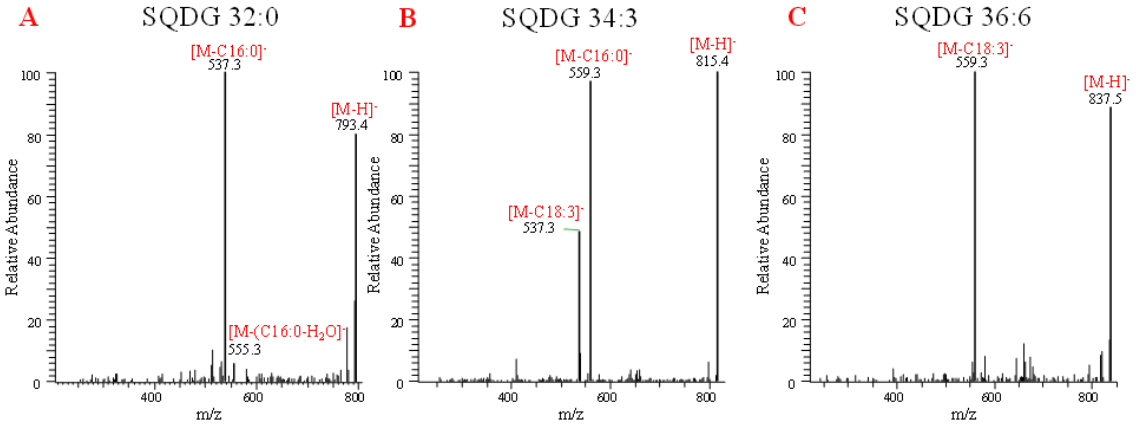


Figure S8. MS/MS spectra of (A) SQDG 32:0, (B) SQDG 34:3, and (C) SQDG 36:6 obtained from the midpoint zone of leaf 3 from inbred B73. Major fragment ions represent the loss of the 16:0 and 18:3 fatty acid side chains.

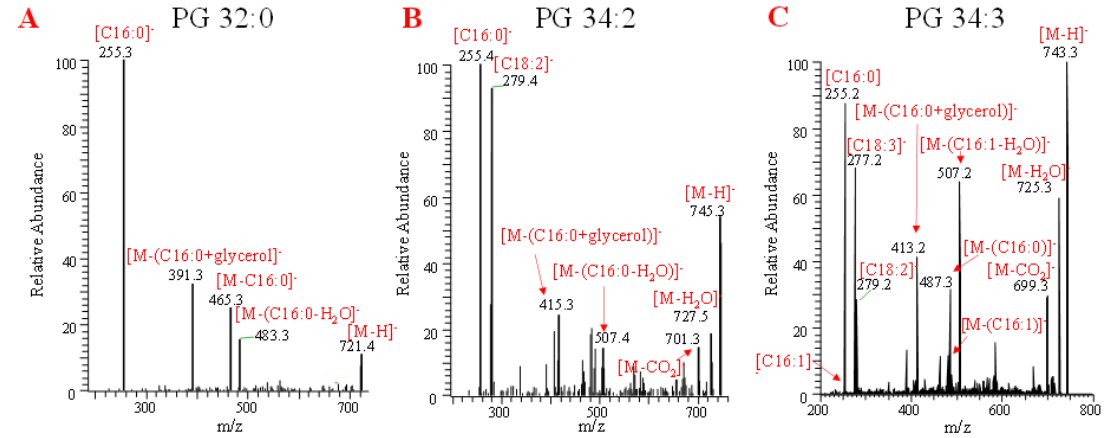


Figure S9. MS/MS spectra of (A) PG 32:0, (B) PG 34:2, and (C) PG 34:3 from the proximal zone of leaf 3 from inbred Mo17, showing two isomers, PG 18:3/16:0 and PG 18:2/16:1. Fatty acid fragments of C18:3 and C18:2 are due to the presence of other phospholipid species (i.e. PE 36:3, PE 36:2) within the isolation mass window used for MS/MS. PGs fragment is predominantly at the sn-2 position, which is the location of 16:0 and 16:1 FAs†. Hence, the loss of C16:0 and C16:1 ($[M - C16:0]^-$, $[M - C16:1]^-$, ($[M - (C16:0\text{-}H_2O)]^-$, $[M - (C16:1\text{-}H_2O)]^-$, $[M - (C16:0 + glycerol)]^-$, and $[M - (C16:1 + glycerol)]^-$) was used for the quantification of PG 18:3/16:0 and PG 18:2/16:1 in Figure 5A.

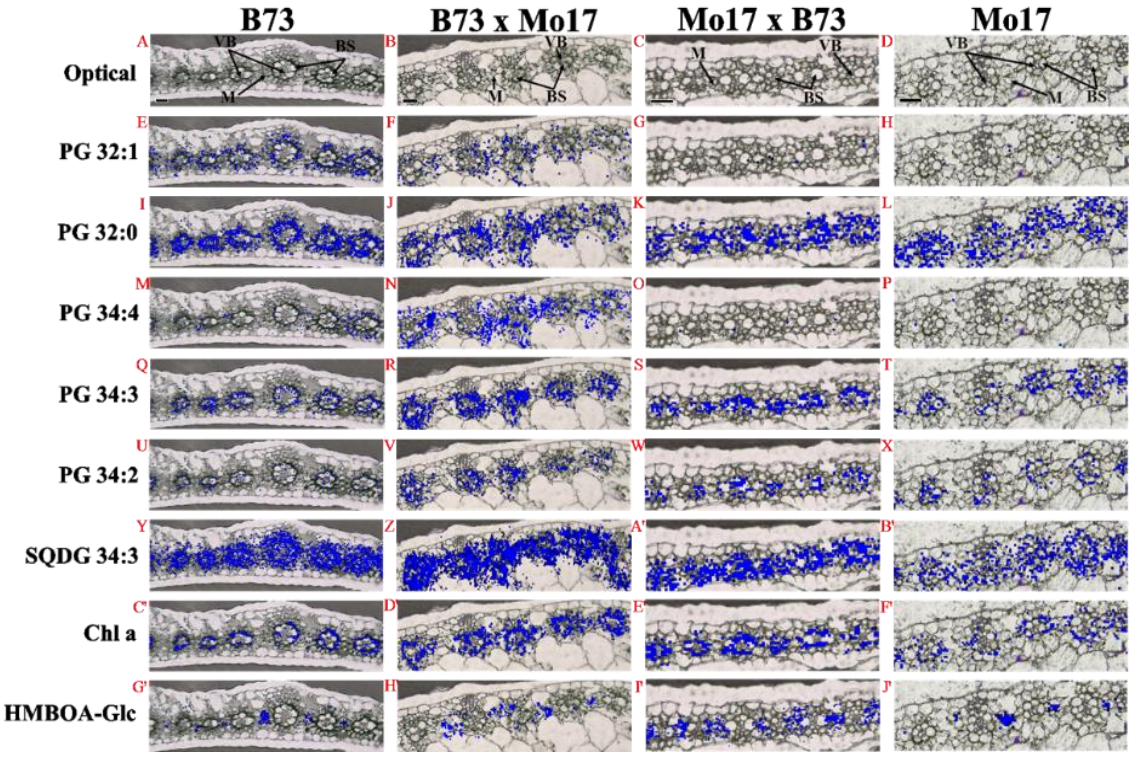


Figure S10. (A-D) Overlay of optical images from the midpoint zone of leaf 3 from B73, B73xMo17, Mo17xB73, and Mo17 with MS images of (E-H) PG 32:1, (I-L) PG 32:0, (M-P) PG 34:4, (Q-T) PG 34:3, (U-X) PG 34:2, (Y-B') SQDG 34:3, (C'-F') Chl a, and (G'-J') HMBOA-Glc. Note that the scale bar is larger for Mo17 and Mo17xB73. Scale bars represent 50 μm.

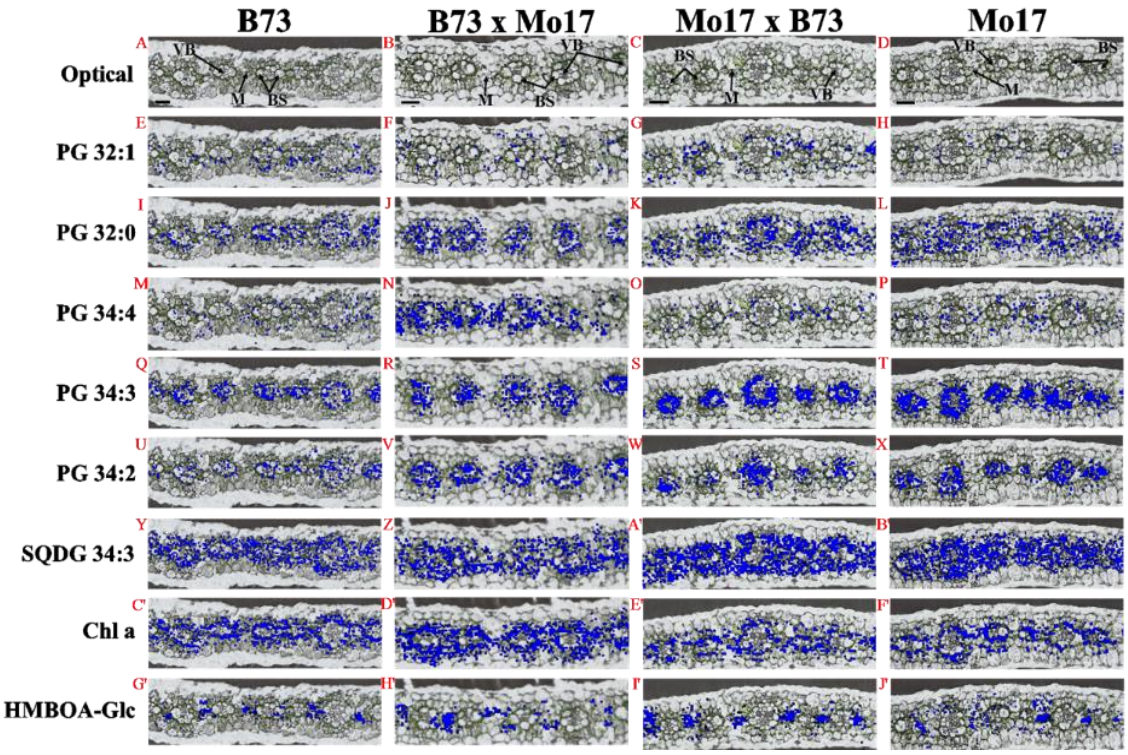


Figure S11. (A-D)Overlay of optical images at the distal cross-section of B73, B73xMo17, Mo17xB73, and Mo17 maize leaves with (E-H) PG 32:1, (I-L) PG 32:0, (M-P) PG 34:4, (Q-T) PG 34:3, (U-X) PG 34:2, (Y-B') SQDG 34:3, (C'-F') Chl a, and (G'-J') HMBOA-Glc. Note that the scale bar is larger for Mo17 and Mo17xB73. Scale bars represent 50 μm

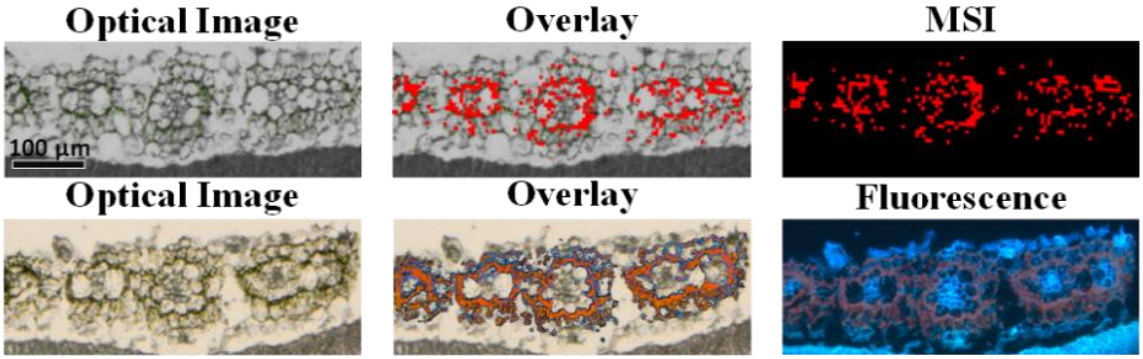


Figure S12. Comparison of chlorophyll *a* localization between (Top) MS and (Bottom) fluorescence images obtained from two consecutive leaf cross-sections. Optical and overlaid images are also shown for the comparison. Considering the slight difference between the two tissues as seen in optical microscope images, MS and fluorescence images are matching quite well. MS and fluorescence images could not be obtained from the same tissue sections. For fluorescence imaging, the tissues needed to be directly attached to glass slides to avoid fluorescence backgrounds. For MS imaging, the tissues needed to be attached with double side tape to maintain tissue integrity during the sample preparation.

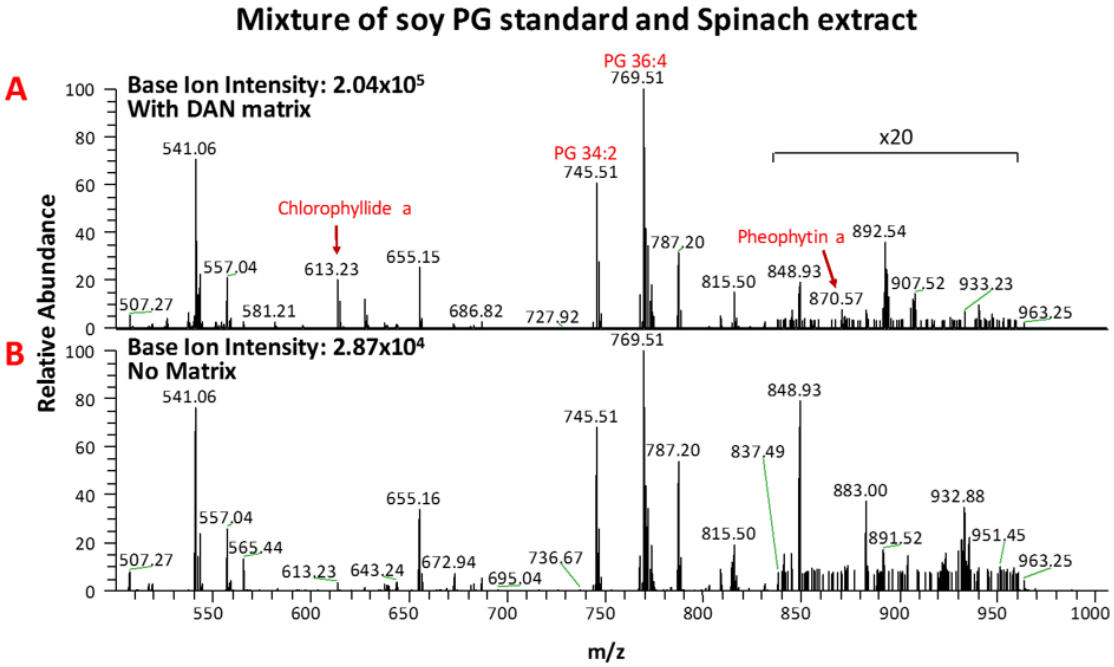


Figure S13. MALDI-MS spectra of a mixture of soy PG standard and spinach extract (A) with and (B) without DAN as the matrix. Spinach extract is used as a source of chlorophyll *a* to study its effect on MALDI ionization. Spinach extract is at the final concentration of 42 mg/mL where the chlorophyll *a* concentration was 0.5 mM according to UV absorption measurement. The final concentration for soy PG standard was 1.25 mM, calculated assuming the average molecular weight of 800 g/mol. DAN concentration was maintained to be low, 0.5 mM, to make it compatible with that of chlorophyll *a*. Base ion signals are $2.04x10^5$ and $2.87x10^4$ ion counts for panel A and B, respectively, demonstrating MALDI efficiency by chlorophyll *a* is only about ~14% even compared with this very low DAN concentration. When compared with typical DAN concentration, ~10mM, MALDI efficiency by chlorophyll *a* was only 1~2%

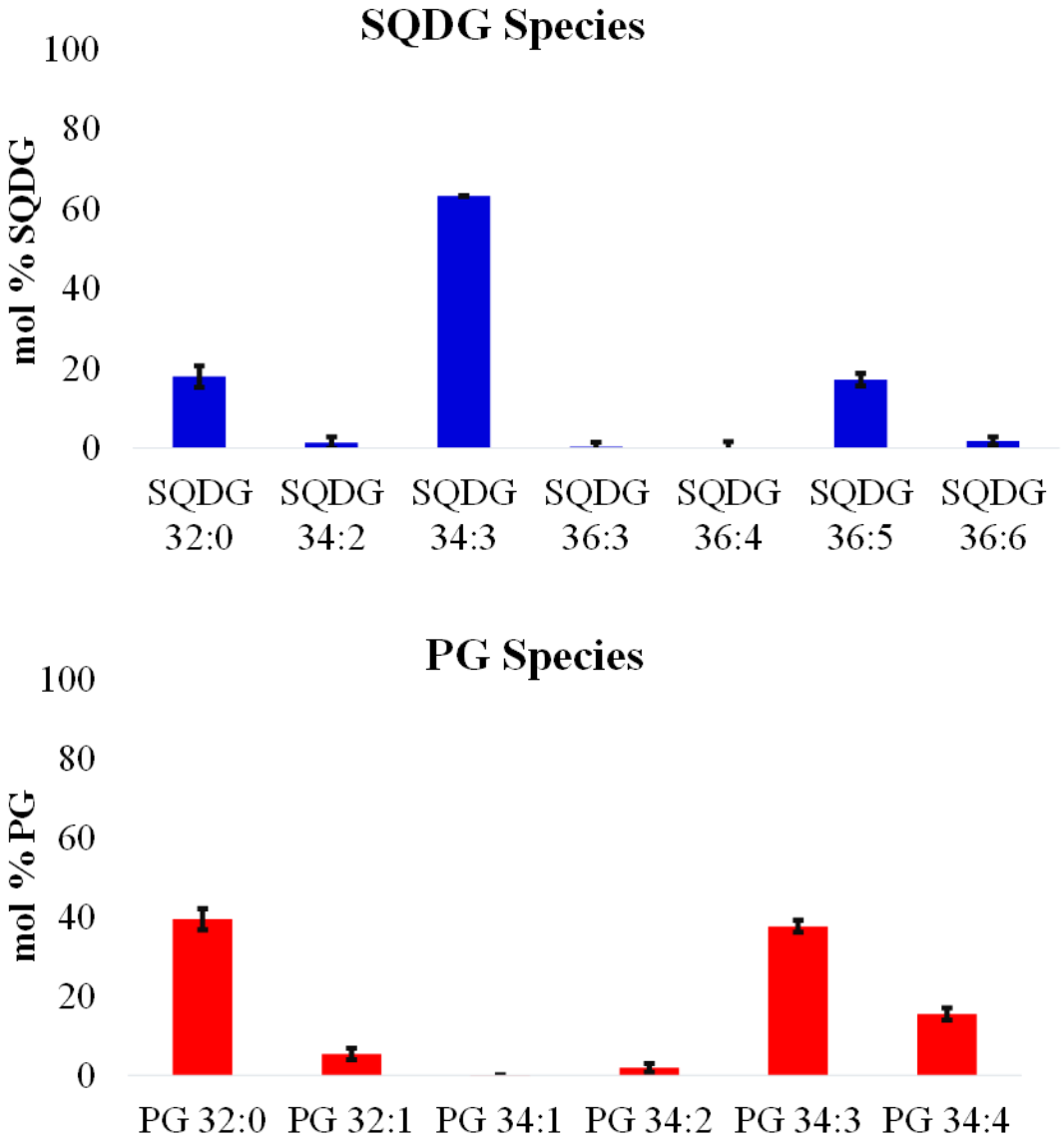


Figure S14. Reproducibility test of MALDI MSI analysis using five consecutive cross-sections of maize leaf. Relative quantification of (Top) SQDG and (Bottom) PG species.

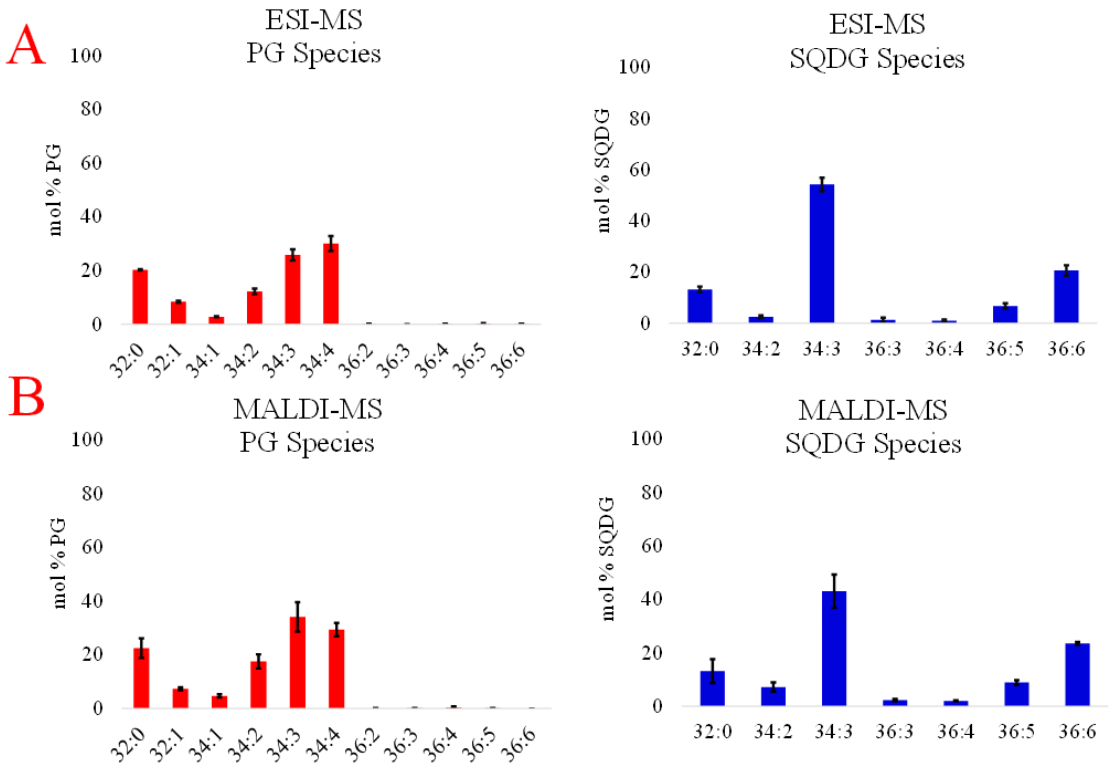


Figure S15. Comparison of PG and SQDG species from total lipid extract of B73 leaf 3 obtained by (A) ESI-FTICR MS and (B) MALDI-MS. Final concentration of leaf extract was 0.025-0.028 mg/ml in ESI-FTICR MS analysis.

Table S1. Quantitative analysis for cell-specific localization of PG species.

	B73				B73xMo17				Mo17xB73				Mo17			
	BS/M			Average	BS/M			Average	BS/M			Average	BS/M			Average
PG 32:1 (**16:0/16:1**)	0.72	0.71	0.66	0.70	0.24	0.23	0.70	0.39	0.01	0.93	0.44	0.46	0.33	0.09	0.78	0.40
PG 32:0 (**16:0/16:0**)	9.34	3.48	2.81	5.21	1.91	2.80	2.26	2.32	0.84	1.40	0.70	0.98	0.71	0.74	1.54	1.00
PG 34:4 (18:3/**16:1**)	0.46	0.90	0.91	0.76	0.48	0.14	0.24	0.29	0.15	0.78	0.44	0.46	0.33	0.22	0.98	0.51
PG 34:3 (18:2/**16:1**)	1.12	2.47	1.41	1.67	1.26	1.25	2.85	1.78	1.24	0.93	1.26	1.14	1.11	1.39	1.32	1.27
PG 34:3 (18:3/**16:0**)	2.47	3.45	2.25	2.72	2.25	2.65	4.80	3.23	2.97	1.61	3.26	2.61	2.25	5.14	2.74	3.38
PG 34:2 (18:2/**16:0**)	82.5	6.46	4.96	31.3	6.34	97.1	11.4	38.3	2.90	26.8	87.0	38.89	57.4	6.91	38.8	34.4

* The numbers represent the ion signal ratios of each PG molecular species in BS vs M cells per unit pixel area obtained from three replicates of midpoint sections. See the experimental section how it was calculated.

* In this analysis, PG ion signals in BS cells are expected to be slightly over-estimated. Because the chloroplasts in BS cells are mostly localized toward the cellular boundary with surrounding M cells, the MS imaging pixels containing both BS and M cells are mostly considered as if they are BS cells. As a result, some PG signals in M cells might have been counted as those in BS cells. Considering the overestimation of BS cells, the color label for the average value of the three replicates is determined as following.

>3	Localization is dominant in BS
2-3	Localization is primarily in BS
0.8-2	Localization is in both M and BS
0.5-0.8	Localization is primarily in M
<0.5	Localization is dominant in M

CHAPTER 4. LIGHT-DEPENDENT CHANGES IN THE SPATIAL LOCALIZATION OF METABOLITES IN *SOLENOSTEMON SCUTELLARIOIDES* (COLEUS HENNA) VISUALIZED BY MATRIX-FREE ATMOSPHERIC PRESSURE ELECTROSPRAY LASER DESORPTION IONIZATION MASS SPECTROMETRY IMAGING

[1]Modified from paper published in Frontiers in Plant Science (2018) 9: 1348

Patrick A. McVey[1,2,3#], Liza E. Alexander[2,3,4#], Xinyu Fu[2,3,4], Bo Xie[2], Katherine-Jo Galayda[1,2,3,5], Basil J. Nikolau[2,3,4] and R.S. Houk[1,2,3]

[#]These authors contributed equally to the work

Abstract

The visualization of foliage color in plants provides immediate insight into some of the compounds that exist in the leaf. However, many non-colored compounds are also present; their cellular distributions are not readily identifiable optically. In this study we evaluate the applicability of mass spectrometry imaging (MSI) via electrospray laser desorption ionization (ELDI) to reveal the spatial distribution of metabolites. ELDI-MSI is a matrix free, atmospheric pressure ionization method that utilizes a UV laser coupled with supplemental ionization by electrospray. We specifically applied ELDI-MSI to determine the spatial distribution of metabolites in Coleus Henna half leaves that were grown with half-sections either fully illuminated

[1]Department of Chemistry, Iowa State University, Ames, IA, USA
[2]Ames Laboratory-US DOE, Ames, IA, USA
[3]Center for Metabolic Biology, Iowa State University, Ames, IA, USA
[4]Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA, USA

or shaded. We monitored dynamic changes in the spatial distribution of metabolites in response to the change of illumination every 7 days for a 28 day period. A novel source-sink relationship was observed between the 2 halves of the experimental leaf. Furthermore, Coleus Henna leaves present visually recognizable sectors associated with the differential accumulation of flavonoids. Thus, we correlated the effect of differential illumination and presence or absence of flavonoids with metabolic changes revealed by the accumulation of carbohydrates, amino acids, and organic acids. The results show the potential of ELDI-MSI to provide spatial information for a variety of plant metabolites with little sample preparation.

Introduction

Traditional mass spectrometry (MS) based methods of metabolite analysis require extraction of metabolites from biological samples, followed by chromatography. Spatial distribution data on metabolites are lost in these procedures. Recently, mass spectrometry imaging (MSI) has made a significant impact in filling this gap and has contributed to more refined understanding of many areas of biology (Amstalden van Hove et al., 2011; Lee et al., 2012; Nilsson et al., 2015; Shroff et al., 2015). Interrogating the location of metabolites in multicellular organisms provides insight into the "sharing" of metabolic processes among different cell types.

Several ionization sources for MSI have been introduced. Each has its own unique advantages and limitations. Matrix assisted laser desorption ionization (MALDI) is the most widely used MSI technique in biology and can obtain spatial resolution as high as ~1 μm (Kompauer et al., 2017; Zavalin et al., 2015). MALDI performed in vacuum, coupled with different applied matrices, allows for the detection of a variety of metabolites (Klein et al., 2015; Shroff et al., 2015). However, the application of the matrix can complicate the use of MALDI especially for metabolites with low *m/z* values (<1000 Da) (Boughton et al., 2016), such as organic

acids, sugars and amino acids. Laser ablation electrospray ionization (LAESI) generally uses an infrared laser to ablate particles from a biological surface, with water in the sample being the main IR absorber. The ablated plant particles undergo post-ionization via interaction with the ESI droplets before entering the mass spectrometer. As generally practiced, LAESI allows ~200 µm lateral resolution, while requiring little sample preparation because there is no applied matrix (Nemes & Vertes, 2007). Similar to the aforementioned methods, matrix assisted laser desorption electrospray ionization (MALDESI) utilizes an applied matrix to absorb radiation from an IR laser coupled with post-ionization via ESI; water ice is an effective matrix for MALDESI (Robichaud et al., 2014).

In the present study, electrospray laser desorption ionization (ELDI) with an ultraviolet laser (wavelength 355 nm) coupled with electrospray (Shiea et al., 2005), was used to analyze the spatial distribution of metabolites in plant samples. In this study, we applied ELDI to Coleus (*Solenostemon scuttellarioides*) leaves that were grown with only half of the leaf surface illuminated. Coleus is a member of the *Labiatae* (mint) family, and different varieties are generally grown as ornamentals. The Coleus variety Henna is known for its serrated foliage. Its upper side is a unique uniform chartreuse to copper color. The underside is dark burgundy. These colors are primarily due to anthocyanin pigments that accumulate in these plants. The color intensity can be affected by illumination, temperature, and other environmental conditions (Lebowitz, 2011).

The present work evaluates ELDI-MSI to study the distribution of anthocyanins, flavonoids, and small metabolites (<1000 Da), including carbohydrates, organic acids and amino acids in Coleus leaves. The effects of different states of illumination during growth are assessed.

Material and Methods

Plant Growth, Maintenance and Light Conditions

Solenostemon scuttellarioides (Coleus Henna spp.) adult plants were obtained from Stam Greenhouse (Oskaloosa, IA). All plants were transferred to LC1 Sunshine Mix soil (Sun Gro Horticulture, Bellevue, WA), watered weekly, and grown in a growth room at 22°C under continuous illumination (2568 Lux or photosynthetic photon flux density 100 µmol of photons m^{-2} sec^{-1}).

For the light regulated experiments, half (through the leaf vein) of each leaf was covered length-wise with aluminum foil (shiny side up). The edges were sealed with scotch tape to ensure the covered region had limited light exposure. The leaves were harvested from the same plant at one-week intervals for four weeks. In all imaging and profiling experiments, three biological replicates (each replicate being images of three plants) were used for MS imaging, optical imaging and metabolic profiling of all metabolites.

Leaves for anthocyanin and non-targeted metabolite profile analysis were harvested each week, cut through the middle vein to separate the light-exposed surface from the shaded/dark-surface, and immediately flash-frozen using liquid nitrogen. The samples were then dried using a vacuum lyophilizer and pulverized using a Mixer Mill 301 (Retsch GmbH, Germany) in 2 mL Eppendorf tubes prior to extraction.

For MS imaging, samples were gently cut from both the foil-covered and uncovered edge of the leaf. Sections measuring approximately 12 mm by 8 mm were immediately analyzed using ELDI. Leaf edges were also visualized under a light microscope.

Tissue Sectioning and Microscopy

For microscopy only, leaf edges were hand-sectioned using a vibratome (TPI-3000; www.tedpella.com) at 70 to 100 µm. The sections were mounted in water and visualized under bright-field using a BH40 compound microscope (Olympus, www.olympus-global.com) equipped with Axio Vision software (Carl Zeiss Inc., Thorwood, NY). Additionally, a fresh leaf was ablated and the depth of the ablation trenches were measured using a Keyence VHX-5000 Digital Microscope.

Anthocyanin Extraction

Compounds were extracted from 5 mg of lyophilized Henna leaf tissue using 300 µL of methanol/ water/ acetic acid (85:15:0.5; v/v/v) and sonicated for 1 hour. Samples were then incubated in dark at 4^{o}C for 2 hours (Wu, Gu, Prior, & McKay, 2004). The samples were centrifuged for 5 minutes at 13,000 rpm and the liquid phase was filtered twice using 13 mm x 0.45 µm Teflon Syringe filters (Supelco, PA) for HPLC MS analysis.

HPLC-MS and MS/MS Analysis of Anthocyanins

The chromatographic separations were performed with an Agilent Technologies 1100 series HPLC. This was coupled with an Agilent Technologies Mass Selective Trap SL detector, equipped with an electrospray ionization (ESI) source with an autosampler/injector and diode array detector (DAD) for LC-MS analysis. A silica-based reverse-phase C18 Atlantis T3 column (2.1 × 150 mm, 3 µm, Waters, Milford, MA) was used for separation. Elution was performed using mobile phases containing 5% formic acid in LC-MS grade water or methanol (MeOH). The flow rate was kept at 0.2 mL/min and the DAD detection was at UV Vis wavelength 520 nm. After 5 µL of sample injection, a gradient was used as follows: 0 to 6 min, 0 to 20% MeOH; 6 to 40 min, 20 to 50% MeOH; 40 to 44 min, 50 to 50% MeOH; 44 to 48 min, 50% to 100% MeOH 48 to 52

min, 100% MeOH; 52-56 min, 100 to 0% MeOH. The column effluent was then introduced via ESI into an Agilent 6210 MSD time-of- flight mass spectrometer operating in positive mode. The ESI capillary voltage was +3.0 kV, nitrogen gas temperature was set to 350°C, drying gas flow rate was 11 L/min, nebulizer gas pressure was 35 psi, skimmer was 65 V, and OCT RF was 250 V. Mass spectra from *m/*z 100 to 2000 were collected and analyzed using Agilent ChemStation Data Analysis. Individual anthocyanin peak areas were generated by Quant Analysis and used to compare different levels of anthocyanins quantitatively. The structure for each anthocyanin was identified from detailed MS/MS analysis performed at various collision energies.

Non-targeted Metabolite Profile Analysis by GC-MS

Metabolite extracts were prepared as described previously (Schmidt et al., 2011). Extracts were prepared from 2 mg of lyophilized leaf tissue. The extracts were spiked with two internal standards: 10 µg of ribitol and 5 µg nonadecanoic acid for polar and non-polar fractions, respectively. To 5 mg of lyophilized tissue, 0.35 ml of hot methanol (60°C) was added. The sample was incubated at the same temperature for 10 min, followed by sonication for 10 min at full power. To this slurry, 0.35 ml of chloroform and 0.3 ml of water were added and the mixture was vortexed for 1 to 3 min. After centrifugation for 5 min at 13,000 g, 200 µl of the upper phase (polar fraction) and 200 µl of the lower phase (non-polar fraction) were separately removed into 2 ml GC-MS vials, and dried in a Speed-Vac concentrator (model SVC 100H, Savant, NY).

The samples were then methoximylated and silylated. For methoximation, 50 µl of 20 mg/mL methoxyamine hydrochloride dissolved in dry pyridine was added. The reaction mixture was shaken at 30°C for 1.5 hours. Silylation was performed by adding 70 µL of *N*, *O*-Bis (trimethyksilyl) trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS) and

incubating at 65°C for 30 min. One microliter of the derivatized samples was injected into the GC-MS in splitless mode. GC-MS analysis was performed using an Agilent 6890 GC interfaced to an Agilent 5973 quadrupole MS with a HP-5ms (5%-Phenyl)-methylpolysiloxane column (30 m x 0.25 mm x 0.25 μm, Agilent). The temperature was programmed from 70 to 320°C at 5°C/min with helium flow rate at 1.0 mL/min and inlet temperature at 280 °C. EI-MS ionization energy was set to 70 eV and the interface temperature was 280°C.

The GC-MS data files were deconvoluted and searched against an in-house MS-library, the NIST 14 Mass Spectral Library using NIST AMDIS software (Stein, 1999) and the Golm Metabolome Database (Hummel et al., 2010). Non-targeted metabolite profiling data obtained by GC-MS analysis, and anthocyanin data obtained from LC-fluorescence are publicly available in the PMR database (http://metnetdb.org/PMR/; Hur et al., 2013). They are available at, metnetweb.gdcb.iastate.edu/PMR/experiments/?expid=279

ELDI-MS

A Waters Synapt G2-S quadrupole time-of-flight mass spectrometer was used for data collection for all MSI. The Waters ESI source was removed and a homemade open-air ESI source (Galayda et al. 2017) was used with samples at ambient pressure. Data were acquired in the mass range from *m/z* 50 to 1200, spectra were summed for 0.3 seconds. The time-of-flight reflectron operated in single-pass mode with a resolution of ~10,000 FWHM for MS images. The TOF was then operated in double-pass mode with a resolution of ~40,000 FWHM, or "high resolution mode", to confirm compound identifications by accurate *m/z* measurements. These confirmatory measurements were done on a different segment taken from the same leaf. Tandem MS quadrupole resolution varied from $\Delta m = 5$ to 12 depending on the analyte with a nominal collision energy of 20 eV. The Synapt was operated using Waters MassLynx V4.1 (SCN851) software.

Data Handling

Spectra were generated from total ion chromatograms (TIC) combined by the MassLynx software. The ".raw" Waters data files were converted to mzML files by Proteowizard Mass Converter Tool. The MzML files were then combined into an imzML file using imzML Converter. This combined image file was then viewed and images were generated from MSiReader V0.06 via the W.M. Keck FTMS Laboratory. All images made within MSiReader had Linear^2 interpolation for image clarity, and used the "Jet" colormap/false color appearance. Co-localization and 3D images were created with MSiReader V1.00. For a few samples in the early stages of this study, images were generated after normalizing individual ion signals to the total ion signal. This normalization did not affect the images noticeably, so normalization was not used for any of the images shown below.

ELDI Compound Identification Protocol

Compounds were identified from the ELDI spectra in a multiple step process to ensure confident assignments. Initial ELDI measurements were done in so-called "sensitivity" mode, m/Δm ~10,000 FWHM. Based on these results, high resolution spectra at ~40,000 were acquired on a different segment of the same leaf for accurate mass measurements. Next, if the compound was of high enough abundance and without background interferences, tandem mass spectra were taken to provide a fragmentation pattern. These high resolution *m/z* values were then put into online metabolite databases (i.e. METLIN) to generate possible compound matches based on accurate mass and tandem MS (if available from both ELDI data and the online database). ELDI data were also compared to the corresponding GC-MS and LC-MS data acquired from Coleus Henna leaves for identity overlap. Every possible ID was given a Δppm value based on the experimental *m/z* value compared with the true *m/z* value of the compound. These Δppm values were generally below

10 ppm. Some compounds were given identifications with greater than 10 Δppm values based on matches with the GC-MS data and/or tandem MS matches with the online database. Some flavonoid peaks were assigned as cyanidin or apigenin derivatives despite no database matches due to highly abundant tandem MS product peaks at *m/z* 287.0585 or 271.0627, respectively. Many METLIN database entries of flavonoids did not have supporting MS/MS data for comparison, with some having only predicted MS/MS spectra. Acquiring tandem MS data with ELDI was successful for most peaks of interest, but was not always possible. It was difficult to acquire tandem MS spectra for highly spatially localized compounds, labeled "localized signal" in subsequent data tables. Compounds with a relatively low abundance did not give a satisfactory fragmentation pattern for identification purposes (labeled "too low abun." in subsequent data tables). Analyte peaks near major background peaks had interferences with their tandem MS spectra. The quadrupole resolution could sometimes be increased to eliminate these interferences, but a subsequent drop in peak intensity sometimes resulted in poor tandem MS results.

ELDI Source

The apparatus has been described (Galayda et al., 2017). Samples were ablated with a Nd:YAG laser (ULTRA, Big Sky Laser Tech, Inc. Bozeman, MT). The third harmonic was used at 355 nm. The laser was operated at a pulse repetition rate of 10 Hz, with a 5 ns pulse width, and an energy of ~250 μJ/pulse (before focusing). This pulse energy was just above the ablation threshold for these samples. The beam was focused onto the sample by a single plano-convex focusing lens (fused-silica, focal length 75 mm), with a nominal spot size of ~125 μm. Samples were cut with a scalpel. Immediately after cutting, the sample segments were mounted on a glass slide using double sided tape. The laser did not penetrate completely through the sample; the underlying tape is not ablated. No matrix was applied, and leaves were pressed lightly to create

an even surface. The optical images shown below were photographed before this pressing step. Thus, the dimensions of the optical images are slightly smaller than those of the MS images in some of the results shown below. Plant samples were then placed 8 mm below the ESI-sample inlet axis on a computer-controlled translation stage (Z825B, Thorlabs, Inc. Newton, NJ). Samples were translated at 0.4 mm/sec beneath the 10 Hz laser beam down the surface of the leaf to insure fresh tissue was constantly being ablated. The distance between the centers of adjacent ablation tracks was 125 µm, providing a lateral resolution of 125 µm. The ablation trench was ~ 30 µm deep. The underlying tape was not ablated. Mass spectra were averaged over 0.3 second intervals. Thus, the ablated volume was 125 µm x 120 µm x 30 µm deep. The leaves were irradiated normal to the sample surface with the laser beam axis ~2 mm downstream from the ESI capillary. The ESI tip was ~10 mm from the sample inlet (Galayda et al., 2017).

A solution of 50% methanol with 0.1% formic acid (99.5% purity, Fisher Scientific) was pumped through a 53 µm ID polyimide coated capillary as the ESI solution (pump: model Z2, Harvard Apparatus, South Natick, MA). Leucine enkephalin was added at 0.1 ppm to the ESI solution for use as a mass calibrant. All data were acquired in positive ESI mode. The ESI voltage for ELDI was +2.5 kV applied to a stainless steel union in the liquid flow line, with the sample cone completing the ESI circuit. The sample inlet was kept at 100^{o} C with a N_2 curtain gas flow of 1 L/hr.

Results

Overview of Experimental Workflow

The mid-rib (Figure 1) was used as the boundary between the two halves of each Coleus leaf. One half of a leaf was covered with aluminum foil to generate leaf tissue that was grown under lower illumination conditions, and the uncovered half of the leaf was used as a control for

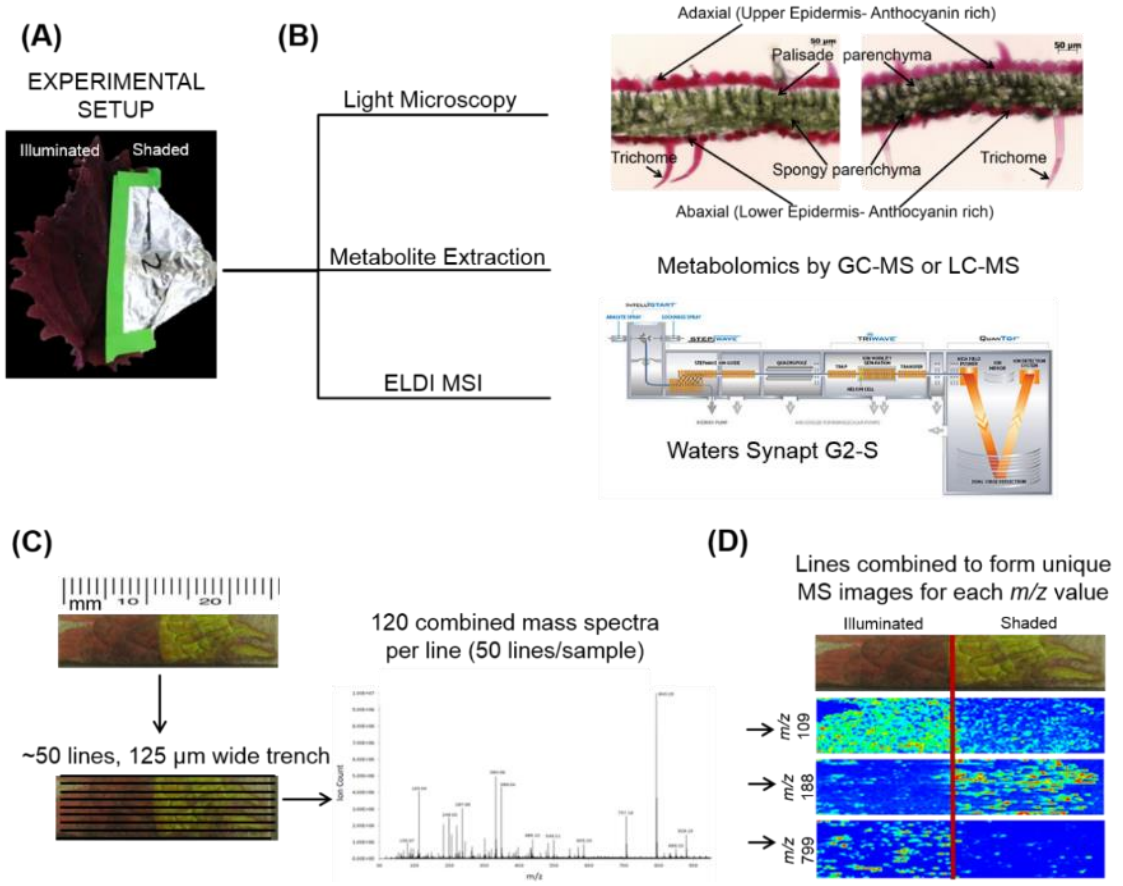


Figure 1. ELDI-MSI experimental workflow. **A)** Individual Coleus leaves were treated for a period of up to 28 days by shading half the leaf and leaving the other half fully illuminated. **B)** Leaf samples were selected for light microscopy, metabolite extraction and metabolomics analysis, and ELDI-MSI analyses with a Waters Synapt G2-S mass spectrometer. **C)** ELDI-MSI data were generated and validated by metabolite profiling of extracts by GC-MS or LC-MS analysis. **D)** Mass spectrometric images were generated for individual ions and aligned with visual images of leaves.

normal levels of illumination. Each leaf was dissected transversely. Segments ~12 mm x 8 mm were collected from both the illuminated and shaded sides of the leaf. These segments were placed on a glass slide and used in ELDI imaging experiments. Metabolite distribution maps were created for 77 chemically identified ions detected by ELDI-MS. Thirteen of these identified ions were observed to decrease in abundance in the shaded side of the leaf, while the abundances of eleven compounds increased in the shaded side. The relative abundances of the remaining 53 ions were unaffected by shading (Tables 1-3). In-parallel, targeted anthocyanin analysis via LC-MS, and

non-targeted global metabolomic analysis via GC-MS was performed on separate leaf samples, which were sampled in triplicate.

Illumination Affects Foliage Color and Anthocyanin Abundances

Under normal illumination conditions both the adaxial (upper) and abaxial (lower) surfaces of the Coleus Henna leaf are a burgundy red-purple color. In the shaded condition, the foliage

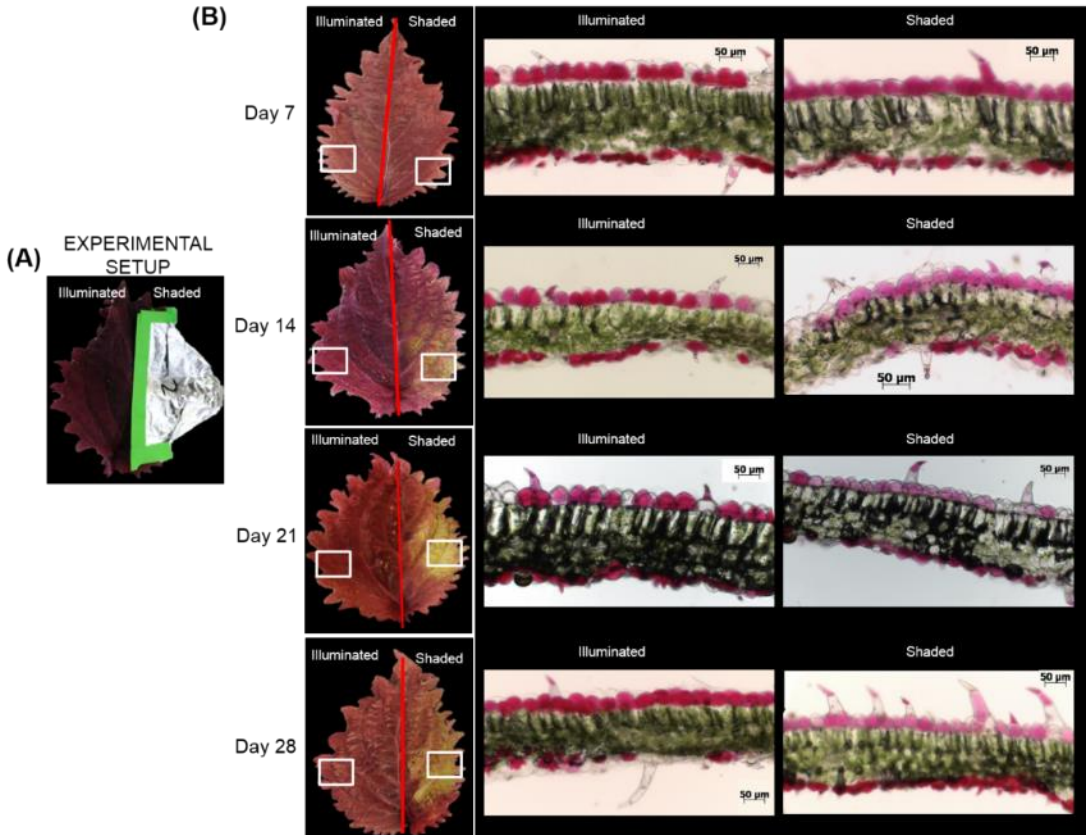


Figure 2. Optical images of Coleus Henna leaves over a 28 day period. **A)** The right-side of each leaf was shaded by wrapping with aluminum foil, whereas the left side was fully illuminated. White boxes denote areas of the leaves that were sectioned for optical microscopic examination in cross section **(B)**. Scale bar = 50 µm

color changes from burgundy red-purple to chartreuse green (Figure 2). Microscopic cross-sections of leaves reveal that the abaxial and adaxial epidermal cells are heavily red-pigmented, and the mesophyll and palisade layers are green pigmented with chlorophylls. Epidermal cells range from 20 to 50 µm deep with cross-sections in the range of 40 to 55 µm long. These cell

dimensions are unaffected by shading. The red pigmentation was decreased in the adaxial epidermis, and unaltered in the abaxial epidermis. This loss of the red pigment from the adaxial epidermis revealed the underlying chlorophyll pigments, imparting the green color to the shaded half of the adaxial surface of the leaf.

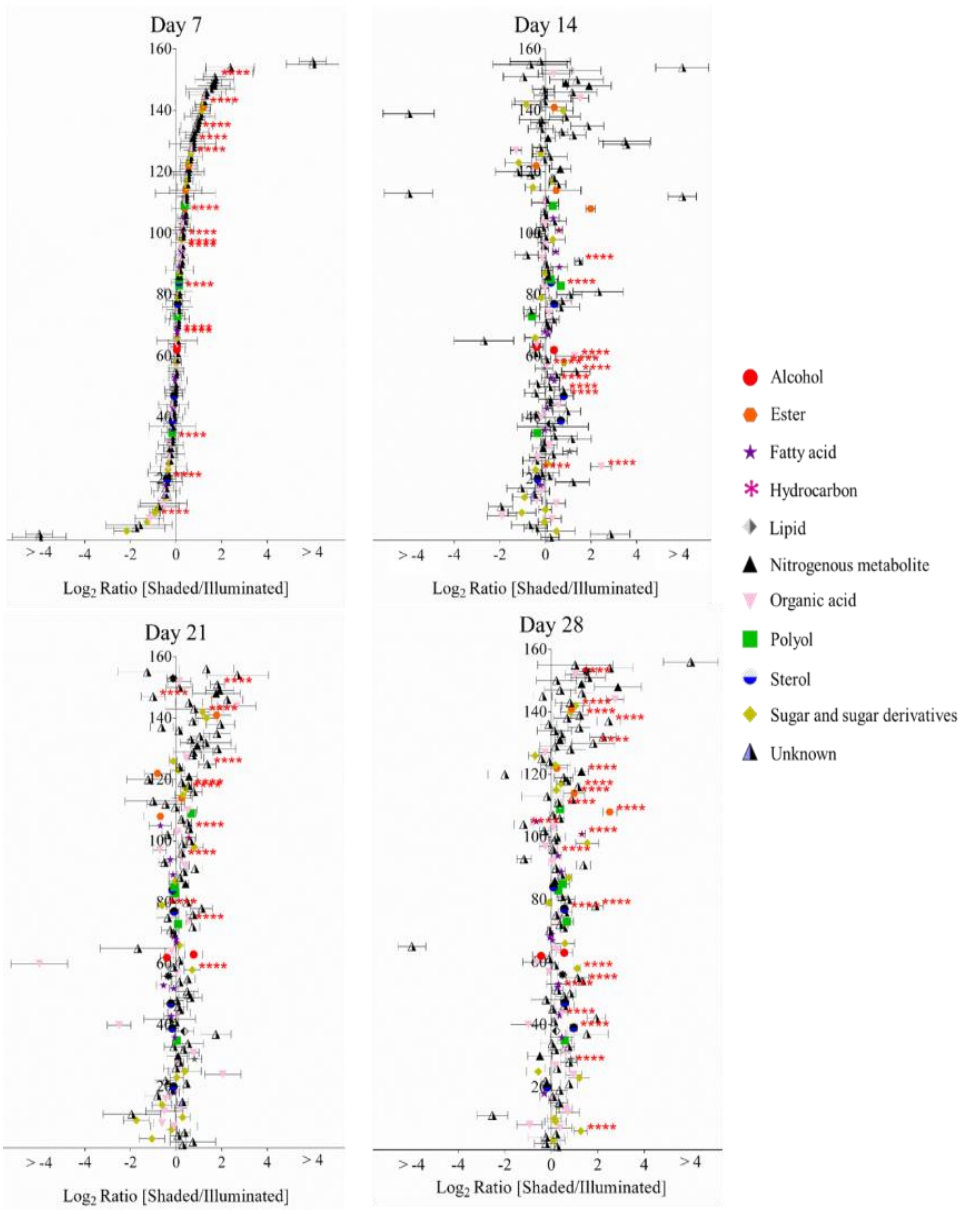


Figure 3. Log-ratio comparison of the differential metabolomes between illuminated and shaded sides of the leaf. The x-axis plots the log-transformed relative abundance ratio of each metabolite in illuminated versus shaded sides. The order of the metabolites (156 analytes, 67 chemically identified) on the y-axis is identical in all four plots, and they are ordered from the lowest to highest value on the x-axis as determined for the 7-day time-point. **** denote p-value <0.05.

Effect of Shading on the Extractable Leaf Metabolome

The GC-MS analyses of metabolite extracts from leaf tissue detected 156 polar and non-polar analytes; 67 of these were identified chemically. These metabolites include alcohols, polyols, sugars, lipids, fatty acids, esters, sterols, hydrocarbons, organic acids, and nitrogenous metabolites. The abundances of most of the chemically identified metabolites remained unaltered irrespective of the illumination conditions (Figure 3). The abundances of a few metabolites, such as gluconic acid, fructose, arachidonic acid, 1-monopalmitin and a hexose sugar, increased in the shaded side of the leaf as time progressed. The abundance of fructose increased in the shaded side of the leaf and became higher than in the illuminated side by the 28th day of the experiment. Benzoic acid showed a unique profile, with abundance decreasing through the time-line of the experiment; its abundance was always higher in the shaded side of the leaf.

Targeted LC-MS analyses measured changes in anthocyanin concentrations during these experiments. Nine cyanidin-based anthocyanins (Figure 4) were identified. Two of these showed differential accumulation between the illuminated and shaded sides of the leaf: cyanidin-

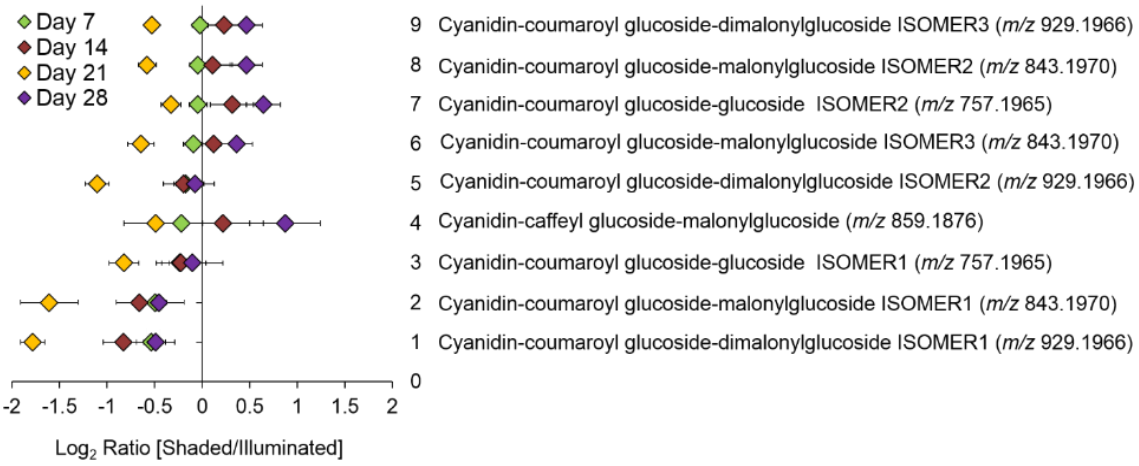


Figure 4. Log-ratio comparison of the differential accumulation of anthocyanins in Coleus half-leaves maintained under full-illumination or shaded for the indicated time-periods. The table names each of the 9 anthocyanins that were identified by LC-MS analysis of extracts prepared from the two leaf halves. The order of the metabolites on the y-axis of the plot is from the lowest to the highest log-ratio values as determined for the 7-day time point.

coumaroylglucoside-malonylglucoside at *m/z* 843.1970 and cyanidin-coumaroylglucoside-dimalonylglucoside at *m/z* 929.1966. The abundance of both these anthocyanins decreased significantly in the shaded side of the leaf after 21 days of the experiment. Day 21 appears to be anomalous compared to the other time points. We suggest that day 21 may represent a metabolic switch, as the leaf is depleted of carbon and subsequently is re-accessed from sink tissues. The accumulation of the other 7 anthocyanins was unaffected by the difference in illumination.

ELDI-MS Identification of Flavonoids

The pigmentation on the abaxial side of the leaf was unaffected by shading, whereas the pigmentation pattern was altered on the adaxial side of the leaf. Therefore, we focused on imaging

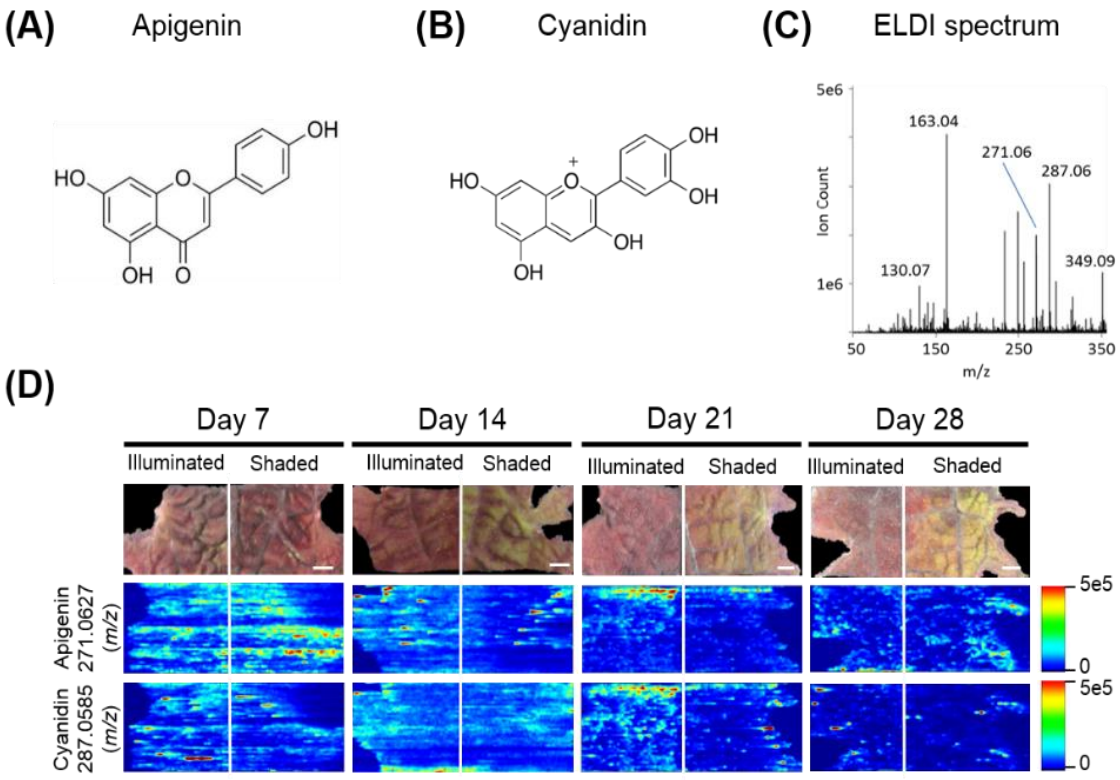


Figure 5. ELDI-MSI analysis of flavonoid backbones. Molecular structure of two aglycone flavonoids: **A)** apigenin; and **B)** cyanidin. **C)** ELDI mass spectrum showing the relative abundance of apigenin ($[M+H]^+$ at *m/z* 271.0627) and cyanidin ($[M]^+$ at *m/z* 287.0585) ions. **D)** Spatial distribution of apigenin and cyanidin in shaded or fully illuminated leaf-halves, at different time-points after initiation of the shading treatment. Scale bar = 2 mm. The MSI intensity scale bar is color coded: red is maximum signal and blue is minimum signal, in this and subsequent figures.

the distribution of the metabolites on the adaxial surface using ELDI-MSI. Many of the ions detected in these ELDI spectra were attributed to flavonoids. Their chemical identity was confirmed by MS/MS experiments that generated fragment ions characteristic of the flavonoid backbones. The identities of these fragment ions were confirmed by matching accurate mass measurements with entries in the Metlin database (Smith et al., 2005). This strategy identified two main classes of flavonoids, based on the aglycone cores: a) apigenin, a flavone **(**Figure 5A**)**, and b) cyanidin, an anthocyanin (Figure 5B).

Apigenin was identified as its protonated ion $[M+H]^+$ at *m/z* 271.0627. All analytes that generated this fragment ion by tandem MS were inferred to be apigenin-based flavones. Cyanidin has a permanent +1 charge and was identified as an $[M]^+$ ion at *m/z* 287.0585. Cyanidin was distinguished from isobaric compounds (e.g., luteolin or kaempferol) based on MS/MS fragments derived from this $[M]^+$ ion (Table 1). Cyanidin-based anthocyanins were identified by cyanidin $[M]^+$ product ions at *m/z* 287.0585. The high intensities of the apigenin-aglycone (*m/z* 271.0627) and cyanidin-aglycone (*m/z* 287.0585) ions in the overall ELDI spectrum (Figure 5C) show that the aglycones have a high relative abundance compared to other low mass molecules observed.

The cyanidin-based anthocyanins are primarily glycosylated and further biochemically modified by malonylation or coumarylation (Table 1). These anthocyanins generated common fragment ions at *m/z* 163.0620 and 147.0485, which were identified as the protonated water-loss ion $[M\text{-}H_2O+H]^+$ of a hexoglycoside, and the protonated water-loss ion $[M\text{-}H_2O+H]^+$ of coumaric acid, respectively. The glycone moiety was not identified directly. However, based on the METLIN and the KEGG pathway databases (Guijas et al., 2018; Smith et al., 2005), and prior characterization of these metabolites in other Coleus lines (Boldt, 2013), these glycosides are subsequently referred to as glucosides. Collectively therefore, we identified 15 cyanidin-based

anthocyanins (Table 1). Similar characterizations identified four apigenin-based flavones, including the apigenin aglycone. Tandem MS generated the apigenin backbone product ion but the specific chemical structures of these apigenin-based flavones were not determined. The ELDI experiments identified five apigenin-based flavones at *m/z* 447.0918, *m/z* 489.1057, *m/z* 619.2228, m/z 635.2041, and *m/z* 743.1454 (Table 1).

Effect of Shading on the Spatial Distribution of Flavonoids

The 15 chemically identified cyanidin-derived anthocyanins can be categorized into three classes. Eight are less abundant in the shaded side of the leaves, one is more abundant in the shaded side, and six are not affected by the difference in illumination (Table 1). This last category includes the cyanidin aglycone, whose abundance are unaffected by shading during the entire 28-day period of the study, same as the apigenin aglycone (Figure 5D). Additional example images are presented in Figure 6, which shows the spatial distribution of five chemically identified cyanidin-based anthocyanins and three apigenin-based flavonoids. These include cyanidin-malonylglucoside (*m/z* 535.1069 $[M]^+$), cyanidin-coumaroyl glucoside (*m/z* 595.1447 $[M]^+$), cyanidin-coumaroylglucoside-glucoside (*m/z* 757.1965 $[M]^+$), cyanidin-coumaroylglucoside-malonylglucoside (*m/z* 843.1970 $[M]^+$), and cyanidin-coumaroylglucoside-dimalonylglucoside (*m/z* 929.1966 $[M]^+$); and the apigenin flavonoids at *m/z* 447.0918 $[M+H]^+$, *m/z* 619.2228 $[M+H]^+$, and *m/z* 743.1454 $[M+H]^+$.

MS images in this paper are obtained by scanning the laser horizontally from left to right. Signals are not acquired between adjacent ablation tracks. This acquisition method generates images that appear to be elongated horizontally, largely due to the way in which the data are acquired. Additional horizontal stretching could occur because of a) there is no delay between pixels to allow recently ablated material to wash out of the ion source, and b) ablation occurs into

room air; there is no controlled gas flow in the ion source. Two examples are the two laterally-elongated features seen in Figure 5D for apigenin, day 14 shaded side, middle of image. The effect is intermittent; it is not always observed. When this elongation is observed, as in Figure 5D, usually multiple adjacent tracks are elongated, which indicate that stretched images are mainly due to the stretched analyte regions in the actual sample.

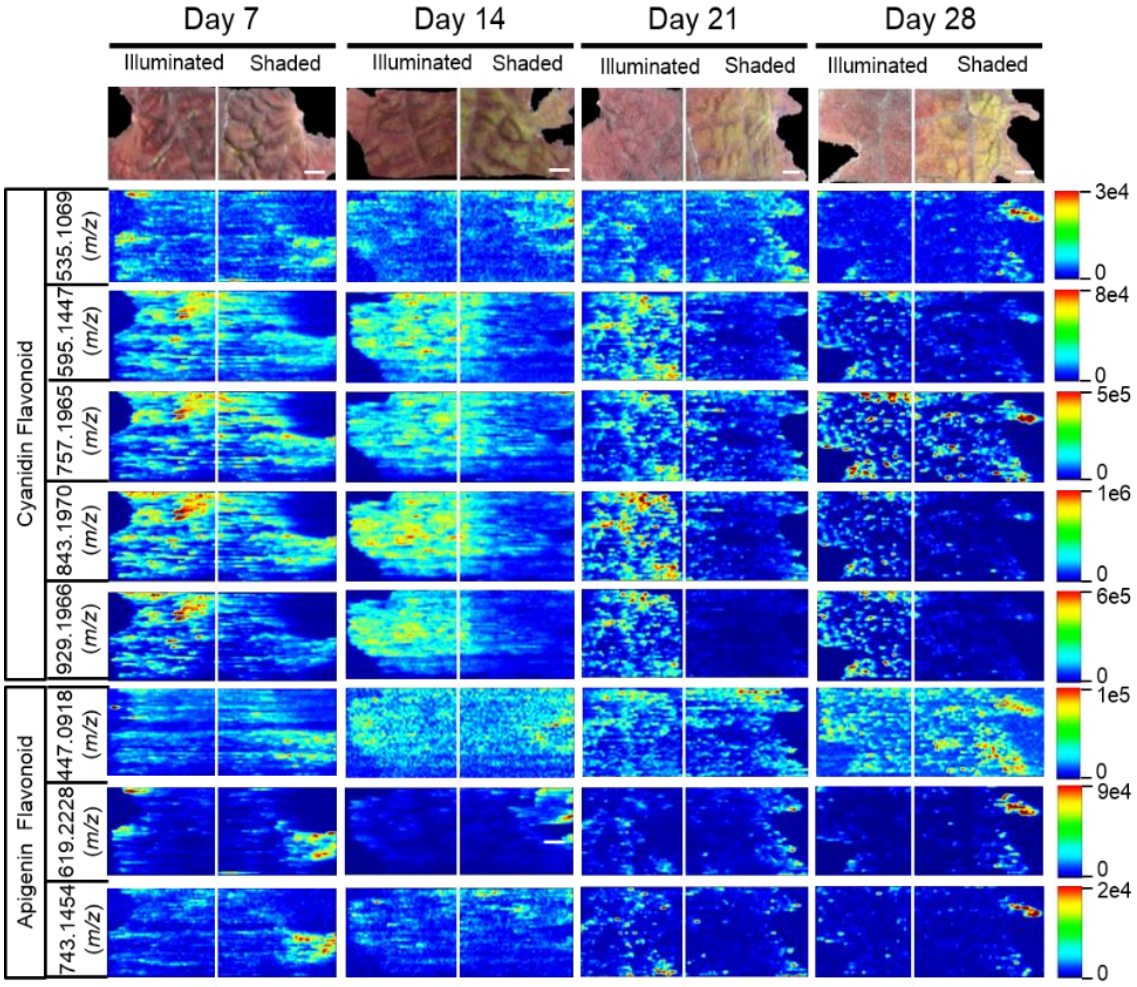


Figure 6. Spatial distribution between shaded and illuminated halves of leaves of cyanidin and apigenin-based flavonoids as affected up to 28 days of shading. The cyanidin-based flavonoids were detected as $[M]^+$ ions, and apigenin-based flavonoids were detected as either $[M+H]^+$ ions. The tentative IDs of the cyanidin-based flavonoids are as follows: cyanidin malonylglucoside (*m/z* 535.1069), cyanidin coumaroylglucoside (*m/z* 595.1447), cyanidin-coumaroylglucoside-glucoside (*m/z* 757.1965), cyanidin-coumaroyl glucoside-malonylglucoside (*m/z* 843.1970), cyanidin-coumaroyl glucoside-dimalonylglucoside (*m/z* 929.1966), and the apigenin-based flavonoids are tentatively identified as apigenin glucuronide (*m/z* 447.0918), apigenin flavonoid (m/z 619.2228), and apigenin flavonoid (*m/z* 743.1454). Scale bar = 2 mm.

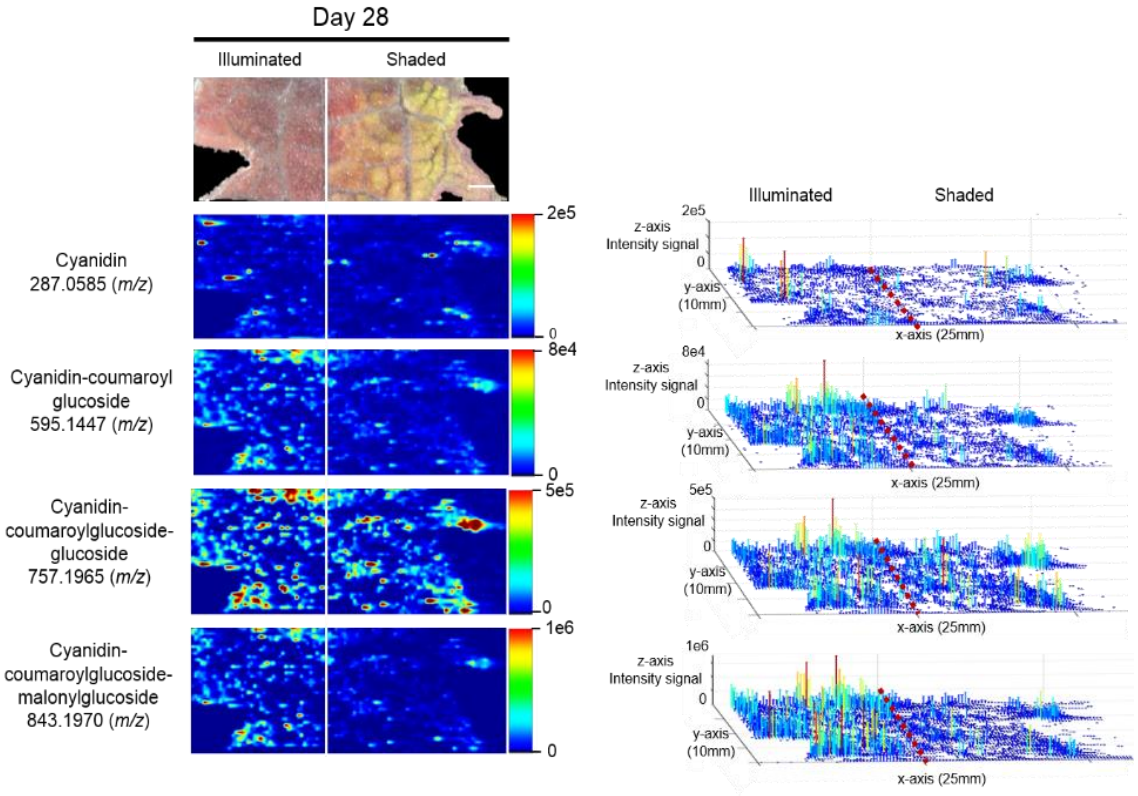


Figure 7. Spatial distributions of cyanidin-coumaroylglucoside-malonylglucoside and its metabolically related species determined by ELDI-MSI analysis of leaf-halves subjected to either full illumination or shaded for 28-days. Right panel represents the 3D representations of the abundance of these metabolically related ions. The x- and y-axes represent spatial coordinates (mm), and the z-axis maps ion intensity. Scale bar = 2 mm.

The anthocyanin with the highest relative abundance, cyanidin-coumaroylglucoside-malonylglucoside (*m/z* 843.1970 $[M]^+$), is initially equally abundant in both illuminated and shaded sides of the leaf. However, starting at 14 days its abundance is reduced more rapidly in the shaded side of the leaf (Figure 6). A similar distribution pattern after 28 days of shading is observed for the structurally related cyanidin- coumaroylglucoside (*m/z* 595.1447 $[M]^+$). In contrast, cyanidin-coumaroylglucoside-glucoside (*m/z* 757.1965 $[M]^+$) is equally abundant in both the illuminated and shaded sides after 28 days (Figure 6). This anthocyanin had the second highest relative abundance and may be the immediate metabolic precursor to cyanidin-coumaroylglucoside-malonylglucoside (Guijas et al., 2018). Three-dimensional temporal distribution patterns of these structurally related anthocyanins (*m/z* 843.1970 $[M]^+$, *m/z* 757.1965

$[M]^+$, *m/z* 595.1447 $[M]^+$ and *m/z* 287.0585 $[M]^+$) after 28-days of shading are presented in Figure 7. Similar to cyanidin-coumaroylglucoside-glucoside, the abundance of the dimalonylated derivative, cyanidin-coumaroylglucoside-dimalonylglucoside (*m/z* 929.1966 $[M]^+$), is lower in the shaded side of the leaf (Figure 6). A precursor to cyanidin-coumaroylglucoside-dimalonylglucoside, the cyanidin-malonylglucoside (*m/z* 535.1069 $[M]^+$) is equally abundant in both the illuminated and shaded sides until 28-days of shading, where its abundance is reduced in the illuminated side of the leaf (Figure 6). All these changes in relative abundance of cyanidin-based anthocyanins, revealed by ELDI-MSI, were confirmed by LC-MS analysis of extracts from these leaves (Figure 4).

The spatial redistribution of two apigenin-based metabolites was revealed by MSI. For example, the apigenin-based flavonoid at *m/z* 619.2228 occurs in both illuminated and shaded sides of the leaves and concentrates in the periphery of the shaded leaf-half (Figure 6). Another apigenin-based flavonoid at *m/z* 743.1454 is initially located in the periphery of the shaded leaf-half. It redistributes temporally to become more evenly dispersed among the two halves of the leaf (Figure 6).

MSI Profiles of Carbohydrates

Twelve sugars and sugar derivatives were chemically identified using three MS-based criteria: a) accurate mass determination with ELDI; b) ELDI-MS/MS fragmentation spectra; and c) integrated retention index and electron impact fragmentation patterns from GC-MS analysis. Most of the 12 sugars were observed as $^{39}K^+$ and/or Na^+ adducts by ELDI (Table 2). Specifically, sucrose was identified by MS/MS fragmentation patterns upon ELDI, and also by GC-MS analysis of extracts. A hexose, possibly glucose, was identified as both Na^+ (*m/z* 203.0560) and $^{39}K^+$ (*m/z* 219.0310) adducts. The identity of a phosphorylated hexose (possibly glucose-phosphate) was

inferred from accurate mass determination of the $^{39}K^{+}$ adduct (*m/z* 312.9720). An organic acid glucoside (possibly coumaroyl glucoside) was identified by MS-/MS experiments as both the Na^{+} (*m/z* 349.0923) and $^{39}K^{+}$ (365.0647) adducts. A heptose was identified by accurate mass determination and by MS/MS fragmentation spectra, but the particular isomer could not be determined.

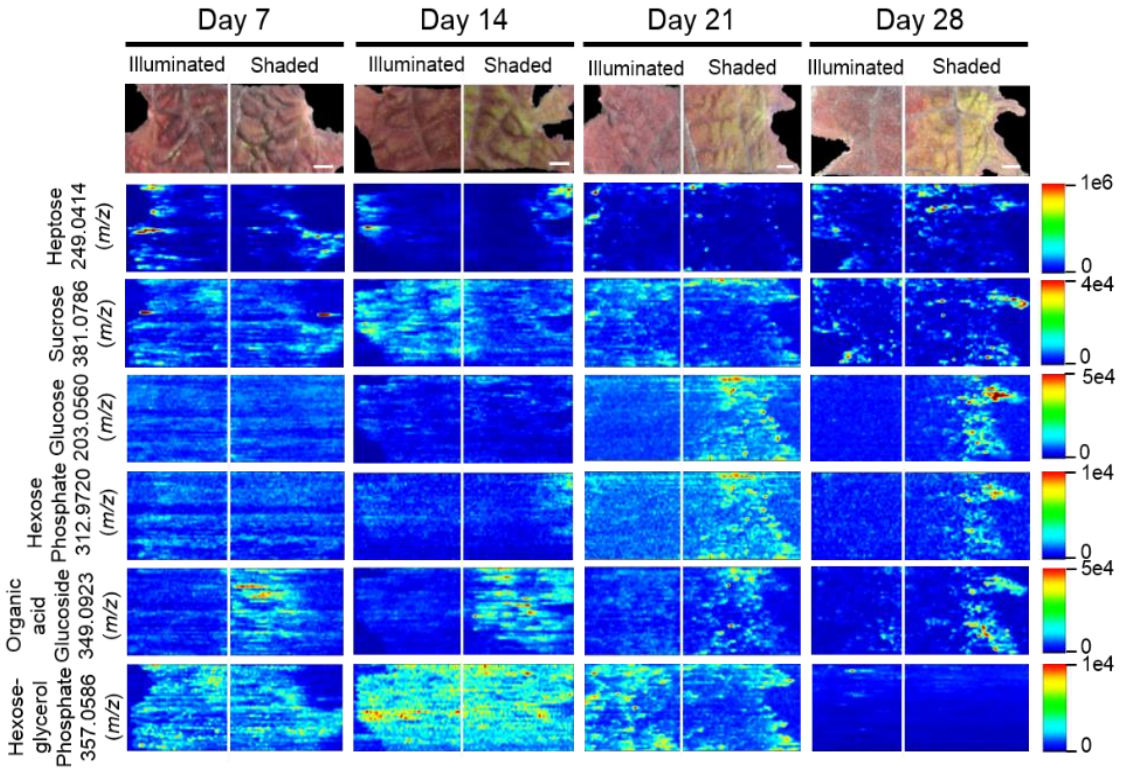


Figure 8. Spatial distributions of sugars determined by ELDI-MSI analysis of leaf-halves subjected to either full illumination or shading for up to 28-days. Each sugar was detected as $[M+Na]^{+}$ or $[M+K]^{+}$ adducts. Scale bar = 2 mm.

The abundances of eight of these sugar metabolites were unaffected by the illumination status of the leaf (Table 2). The spatial distributions of six typical metabolites are shown in Figure 8. Three of these sugar metabolites (*m/z* 203.0560, *m/z* 312.9720, *m/z* 349.0923) become more abundant over time in the shaded side of the leaf. Finally, a sugar-derivative, believed to be hexose-glycerol phosphate (*m/z* 357.0586), as determined by accurate mass, became less abundant across the entire leaf after 28 days of shading (Figure 8).

MSI Profiles of Organic and Amino Acids

Table 3 lists the carboxylic acids and amino acids that were observed by ELDI-MSI. Seven of these metabolites were identified as H^+ adducts, 15 as H^+ adducts accompanied by water loss, two as Na^+ adducts, and three as $^{39}K^+$ adducts. Several of these compounds were observed as multiple adducts; individual molecules of a given compound had one of either H^+, Na^+ or $^{39}K^+$ attached in the same spectrum.

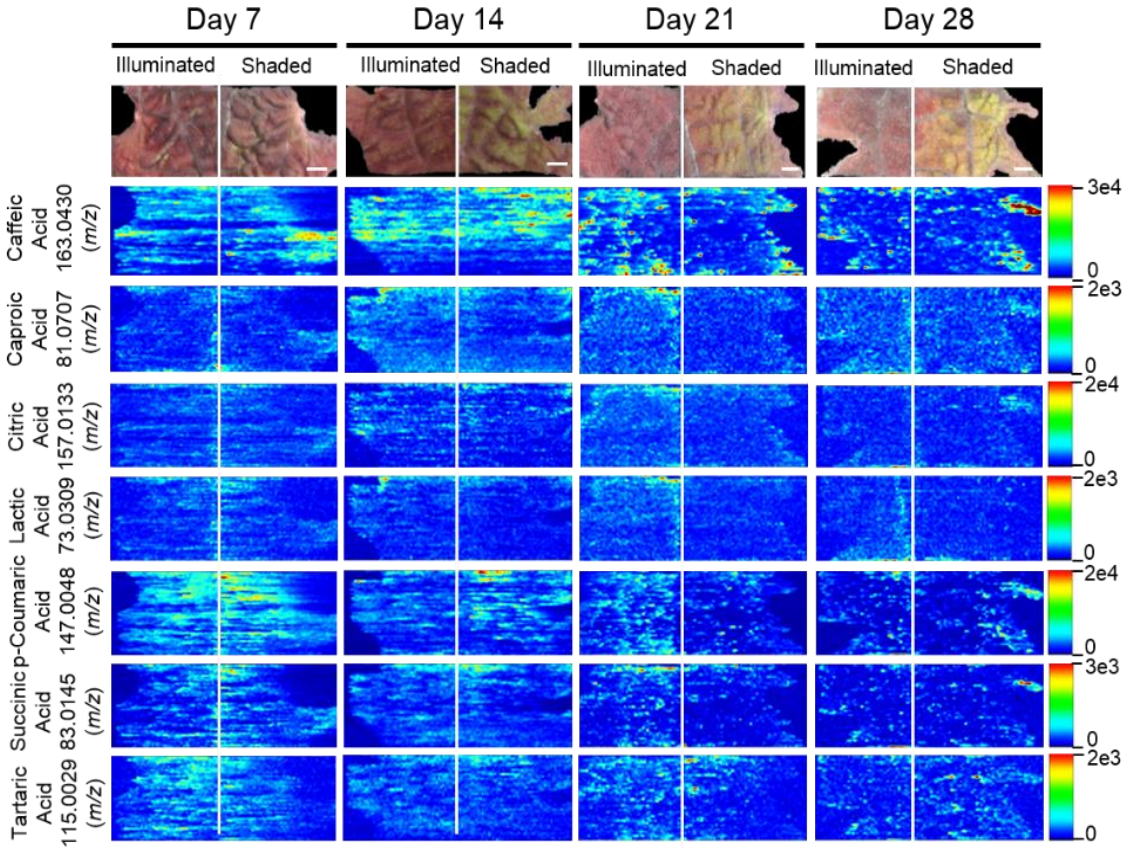


Figure 9. Spatial distributions of organic acids determined by ELDI-MSI analysis of leaf-halves subjected to either full illumination or shading for up to 28-days. Scale bars = 2 mm.

Several analytical strategies were integrated to confirm the chemical identity of these organic acids, including accurate mass determination, MS/MS experiments, and GC-MS analysis of derivatized metabolite extracts. The latter strategy was also used to confirm the relative concentrations between the illuminated and shaded sides of the leaves. Collectively these analyses

identified 23 organic acids. Eight were identified with ELDI-MS and the molecular images of these metabolites (Figure 9) indicate their abundances were unaffected by the shading of the leaf, except for glycolic acid seen only by GC-MS.

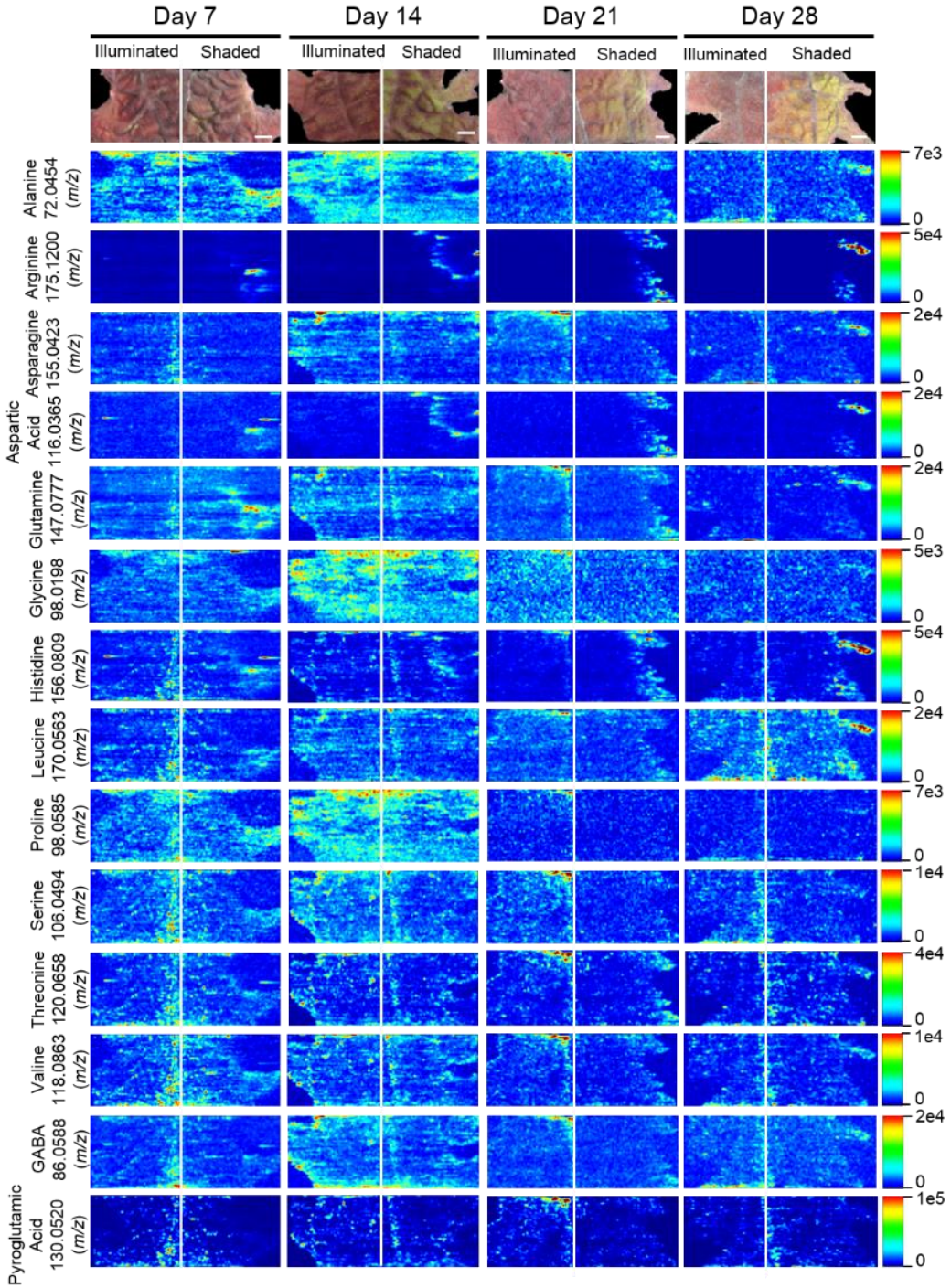


Figure 10. Spatial distribution of amino acids determined by ELDI-MSI analysis of leaf-halves subjected to either full illumination or shading for up to 28-days. Scale bar = 2 mm.

The locations of 12 proteogenic amino acids (alanine, arginine, asparagine, aspartic acid, glutamine, glycine, histidine, leucine, proline, serine, threonine, and valine) were determined by ELDI-MSI. The non-proteogenic amino acids were pyroglutamic acid and GABA (Figure 10). Arginine, aspartic acid, and histidine were considerably more abundant in the shaded side of the leaf, although in the case of histidine this increase occurred after the half-leaf was shaded for 21 days. The spatial distribution of seven amino acids were unaffected by shading, but the relative abundance of four amino acids changed as time progressed. Specifically, leucine abundance increased at 28 days, whereas the abundance of alanine (decreased from day 14 to 28), glycine (peaked day 14 decreased from day 21 onwards), and proline (depleted by day 21) was reduced during this time period.

Discussion

Several MSI methods have been utilized to obtain spatial distribution data of metabolites in plant samples. Each offers unique advantages and drawbacks (Boughton et al., 2016; Lee et al., 2012). The present work demonstrates the potential of applying ELDI for MSI. ELDI is analogous to LAESI, with the exception that ELDI uses ultraviolet radiation for ablation (e.g., 355 nm laser), whereas LAESI generally uses infrared radiation, for example a 2940 nm mid-IR laser (e.g., Nemes & Vertes, 2007). Although ELDI has previously been reported with fungi (i.e., *Ganoderma lucidum* and *Antrodia camphorate*) (Huang et al., 2012), there is only one other moderately extensive plant application study using ELDI, LAESI, or MALDESI (Etalo et al., 2015). The spatial resolution was limited to 500 µm in that work (Etalo et al., 2015). Many proof-of-concept experiments using LAESI-MSI on plants have been reported (Bartels & Svatoš, 2015).

Because plants accumulate large quantities of UV-absorbing molecules (e.g., flavonoids, chlorophylls, terpenes, phenolics) the plant leaf itself serves as a pseudo-matrix. Ablation of the

plant pseudo-matrix enables ELDI-based observation of additional molecules that are not UV absorbers (e.g., sugars, amino acids and organic acids). In both ELDI and LAESI methods, particles undergo post-ionization via interaction with ESI droplets, and the resulting ions are then extracted into the mass spectrometer. Although LDI without post-ionization via ESI can achieve much higher spatial resolution, issues associated with ionization transfer still need to be addressed (Hölscher et al., 2009).

One concern about our ELDI method is possible spatial heterogeneity of the UV-absorbing pseudo-matrix. Coupling electrospray and ablation with a UV laser appears to minimize these issues. Both UV and non-UV absorbing compounds can be measured at atmospheric pressure. For example, Figure 11 illustrates the localization of the UV-absorbing flavonoid cyanidin-

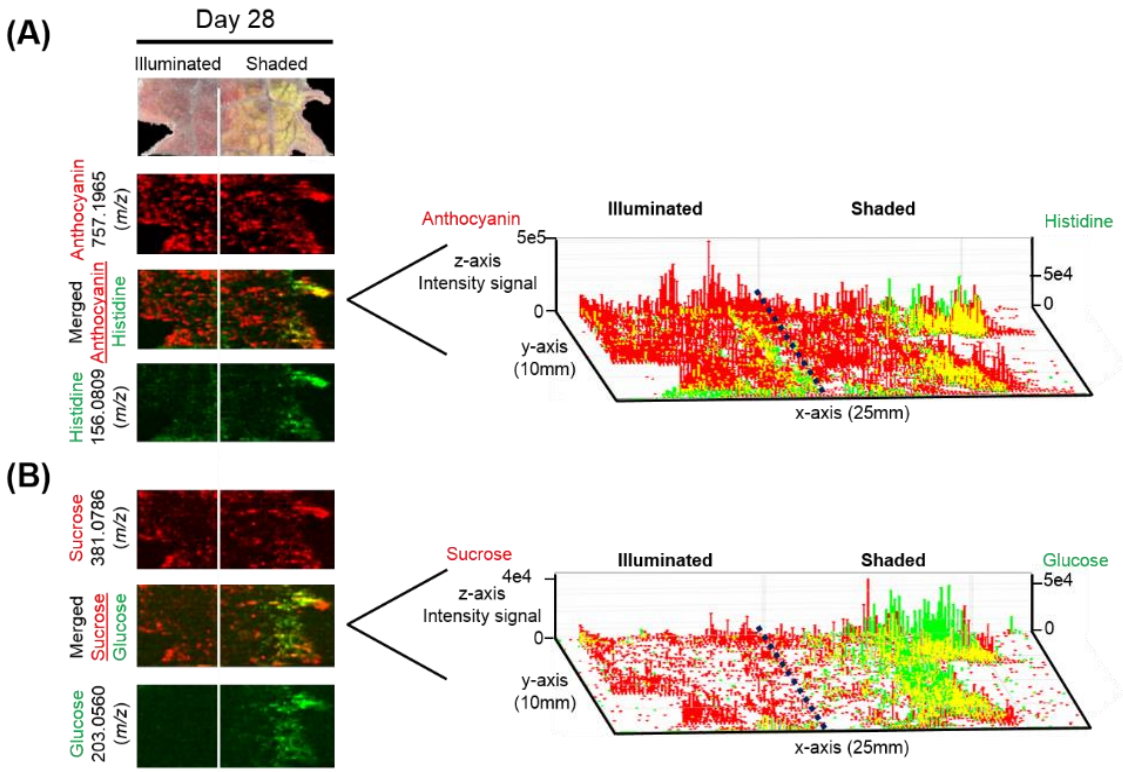


Figure 11. 3D representation of the localization of UV-absorbing and non-UV absorbing metabolites. Distributions of **A)** anthocyanin 757.1965 (*m/z*) (red) along with histidine 156.0809 (*m/z*) (green) and **B)** sucrose 381.0786 (*m/z*) (red) along with glucose 203.0560 (*m/z*) (green) are shown. Scale bar = 2 mm. Right panel represents the 3D representations of the abundances of these metabolically related ions. The x- and y-axes represent spatial coordinates (mm), and the z-axis maps ion intensity.

coumaroylglucoside-glucoside (*m/z* 757.1965 [M^+]) and the localization of the non-UV absorbing metabolites, histidine, sucrose and glucose. Non-absorbing analytes can be observed readily in regions where the abundances of the pseudo-matrix flavonoids are relatively low (Figure 11). These comparisons indicate that ELDI-based MSI can be used to localize the distribution of non-UV absorbing metabolites despite large spatial changes in concentration of flavonoids or other possible pseudo-matrix flavonoids are relatively low (Figure 11). These comparisons indicate that ELDI-based MSI can be used to localize the distribution of non-UV absorbing metabolites despite large spatial changes in concentration of flavonoids or other possible pseudo-matrix compounds.

Coleus is convenient because its leaves are sectored visually. This reflects the spatial arrangement of the underlying flavonoids (Boldt, 2013; Nguyen & Cin, 2009; Nguyen et al., 2008). Moreover, in Coleus the sectoring of the adaxial surface is different from that of the abaxial surface. The adaxial surface is further modifiable by changes in the exposure to illumination (Nguyen & Cin, 2009; Logan et al., 2015). In addition, the depth of the leaf adaxial cells matches the penetration depth of the laser (~30 μm). Thus, the images presented herein reflect the abundance of metabolites in the epidermal cell layer of the leaf.

Correlated with the visible changes in the pigments of the leaf in response to shading, major metabolic changes in anthocyanins were visualized by ELDI-MS and quantitatively confirmed by LC-MS analysis of metabolite extracts. The cyanidin glycosides are the most abundant anthocyanins in the illuminated side of the leaves, which correlates with their photo-inhibitory potential (Steyn et al., 2002). Consistent with this photo-inhibitory attribute, the abundances of 17 of the 39 detected anthocyanins decreased upon shading of the leaf. Moreover, in response the illumination status of the leaf, there is a coordinated change in the differential abundance of the structurally related anthocyanins, cyanidin-coumaroylglucoside-glucoside, cyanidin-malonyl

glucoside, and the aglycone cyanidin. The ability to visualize this coordination at a spatial level provides additional insights to their potential metabolic interconnections. Namely, the enhanced accumulation of cyanidin-coumaroylglucoside-malonylglucoside, in response to illumination, may be associated with increased biosynthesis. The correlated changes in the levels of the potential precursors (i.e., cyanidin-malonyl glucoside) (Figure 7) support this hypothesis.

Multi-cellular photosynthetic plants are characterized by a series of source-sink tissues, that share the metabolic tasks of converting inorganic precursors, such as CO_2, ammonia, phosphate etc., to organic constituents that are normally stored in sink tissues, such as seeds, tubers etc. (Basu et al., 1999; Wardlaw, 2006; Lemoine et al., 2013; Osorio et al., 2014; Paul & Foyer, 2001; Roitsch, 1999; Turgeon, 1989). Source-sink relationships can be genetically programmed and are further modified by environmental abiotic or biotic stimuli that are mediated by small molecules (e.g., sugars and amino acids) (Krapp & Stitt, 1995; Lemoine et al., 2013; McCormick et al., 2008; Paul & Driscoll, 1997; Roitsch, 1999). ELDI-MSI can image metabolic changes associated with the induction of an artificial new source-sink relationship between two halves of a leaf. The fully illuminated half of the leaf serves as the source tissue that fixes carbon, primarily in the form of sugars. These compounds are exported to the shaded half of the leaf that serves as the new sink tissue, whose strength increases with increasing time of shading (Figure 8). (Bagnall et al., 1988; Islam et al., 2005; Paul et al., 1992; Steyn et al., 2002).

ELDI revealed that the distribution of sugars (tetroses, pentoses, heptoses, and sucrose) were not significantly affected by the difference in the illumination between the two halves of the leaves. Only the distributions of glucose and a hexose phosphate (presumably glucose-6-phosphate) were affected; they increased on the shaded side (Figure 8). Therefore, the illuminated source-side of the leaf compensated for the reduced photosynthesis that was imposed by shading

the other half of the leaf. This sugar-based interrelationship between source-sink tissues often manifests coordinated changes in amino acid metabolism, associated with the affiliation between carbon and nitrogen metabolism (McCormick et al., 2006; Paul & Driscoll, 1997). These changes in amino acid metabolism are usually associated with photosynthetic source tissues, where changes in RUBISCO levels (the major sink for amino acids) can drastically affect free amino acid pools (Nielsen et al., 2002; McCormick et al., 2008; Paul & Driscoll, 1997). Thus, in our studies we visualized increasing levels of Arg, Asp, His in the shaded side of the leaf (Figure 10), probably reflecting the turnover of RUBISCO as photosynthetic capacity was reduced in the dark, sink-side of the leaf.

Conclusion

This study demonstrates the capabilities of ELDI-MSI for identification and spatial characterization of a wide variety of compounds in plant tissues with minimal sample preparation requirement. The ability to generate spatial distribution data that are consistent with biological explanations provides confidence in the validity of the observations. Future experiments include a) separation of isobaric ions by ion mobility, and b) implementation of procedures for quantification with spatial resolution by ELDI.

Compared to MALDI or other methods that require an added matrix, ELDI has the advantage of less sample preparation; however its sensitivity and spatial resolution are poorer. Much of the plant sample remains intact after ELDI analysis. Compared to LAESI with the commercial source, their spot size and spatial resolution are ~200 μm, while ELDI as shown here is ~125 μm. All plants that we have analyzed absorb at 355 nm well enough for laser ablation analysis. While the classes of compounds we report here were readily observed, other classes, such as lipids and peptides, were difficult to observe in our ELDI experiments.

Acknowledgements

This research was supported by funds from the U.S. Department of Energy, Office of Biological and Environmental Research through award DESC0014038 to Iowa State University. B.J.N. and L.E.A. were partially supported by the Center for Metabolic Biology, Iowa State University (www.metabolicbiology.iastate.edu). The ELDI-MSI instrumentation was funded by the Office of Basic Energy Sciences, Separations and Analysis Program, U. S. Department of Energy. Metabolomics analyses were conducted at the Iowa State University's W. M. Keck Metabolomics Research Laboratory (metabolomics.biotech.iastate.edu). We thank Dr. Lucas J. Showman and Dr. Ann M. Perera for their expert advice. Microscopic data were gathered at Iowa State University's Microscopy and NanoImaging Facility, and we thank Ms. Tracey Stewart and Dr. Harry T. Horner for their expert advice and assistance.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Amstalden van Hove, E. R., Smith, D. F., Fornai, L., Glunde, K., & Heeren, R. M. A. (2011). An Alternative Paper Based Tissue Washing Method for Mass Spectrometry Imaging: Localized Washing and Fragile Tissue Analysis. *Journal of the American Society for Mass Spectrometry*, *22*(10), 1885–1890. http://doi.org/10.1007/s13361-011-0203-z

Angel, P. M., & Caprioli, R. M. (2013). Matrix-Assisted Laser Desorption Ionization Imaging Mass Spectrometry: In Situ Molecular Mapping. *Biochemistry*, *52*(22), 10.1021/bi301519p. http://doi.org/10.1021/bi301519p

Bagnall, D. J., King, R. W., & Farquhar, G. D. (1988). Temperature-dependent feedback inhibition of photosynthesis in peanut. *Planta (Germany, F.R.)* 175:348. https://doi.org/10.1007/BF00396340

Bartels, B., & Svatoš, A. (2015). Spatially resolved in vivo plant metabolomics by laser ablation-based mass spectrometry imaging (MSI) techniques: LDI-MSI and LAESI. *Frontiers in Plant Science*, *6*, 471. http://doi.org/10.3389/fpls.2015.00471

Basu, P. S., Sharma, A., Garg, I. D., & Sukumaran, N. P. (1999). Tuber sink modifies photosynthetic response in potato under water stress1Publication No. 1408, CPRI, Shimla.1. *Environmental and Experimental Botany*, *42*(1), 25–39. https://doi.org/10.1016/S0098-8472(99)00017-9

Boldt, J. K. (2013). Foliar anthocyanins in coleus and ornamental grasses: accumulation, localization, and function. Retrieved from the University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/150590.

Boughton, B. A., Thinagaran, D., Sarabia, D., Bacic, A., & Roessner, U. (2016). Mass spectrometry imaging for plant biology: a review. *Phytochemistry Reviews*, *15*(3), 445–488. http://doi.org/10.1007/s11101-015-9440-2

Etalo, D. W., De Vos, R. C. H., Joosten, M. H. A. J., Hall, R. D. (2015). Spatially Resolved Plant Metabolomics: Some Potentials and Limitations of Laser-Ablation Electrospray Ionization Mass Spectrometry Metabolite Imaging. *Plant Physiology,* 169(3), 1424-1435. https://doi.org/10.1104/pp.15.01176

Galayda, K. (2017). Use of ultraviolet laser ablation electrospray ionization mass spectrometry for plant metabolite imaging. *Graduate Theses and Dissertations. 15517.* Retrieved from https://lib.dr.iastate.edu/etd/15517

Guijas, C., Montenegro-Burke, J. R., Domingo-Almenara, X., Palermo, A., Warth, B., Hermann, G., Koellensperger, G., Huan, T., Uritboonthai, W., Aisporna, A. E., Wolan, D. W., Spilker, M. E., Benton, H. P., & Siuzdak, G. (2018). METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Analytical Chemistry*, *90*(5), 3156–3164. http://doi.org/10.1021/acs.analchem.7b04424

Hölscher, D. , Shroff, R. , Knop, K. , Gottschaldt, M. , Crecelius, A. , Schneider, B. , Heckel, D. G., Schubert, U. S. and Svatoš, A. (2009). Matrix-free UV-laser desorption/ionization (LDI) mass spectrometric imaging at the single-cell level: distribution of secondary metabolites of Arabidopsis thaliana and Hypericum species. *The Plant Journal*, *60*(5), 907–918. http://doi.org/10.1111/j.1365-313X.2009.04012.x

Huang, M.-Z., Cheng, S.-C., Jhang, S.-S., Chou, C.-C., Cheng, C.-N., Shiea, J., Popov, I.A., Nikolaev, E. N. (2012). Ambient molecular imaging of dry fungus surface by electrospray laser desorption ionization mass spectrometry. *International Journal of Mass Spectrometry*, *325–327*, 172–182. https://doi.org/10.1016/j.ijms.2012.06.015

Hummel, J., Strehmel, N., Selbig, J., Walther, D., & Kopka, J. (2010). Decision tree supported substructure prediction of metabolites from GC-MS profiles. *Metabolomics*, *6*(2), 322–333. http://doi.org/10.1007/s11306-010-0198-7

Hur, M., Campbell, A. A., Almeida-de-Macedo, M., Li, L., Ransom, N., Jose, A., Crispin, M., Nikolau, B. J., Wurtele, E. S. (2013). A global approach to analysis and interpretation of

metabolic data for plant natural product discovery. *Natural Product Reports*, *30*(4), 565–583. http://doi.org/10.1039/c3np20111b

Islam, M. S., Jalaluddin, M., Garner, J. O., Yoshimoto, M., & Yamakawa, O. (2005). Artificial shading and temperature influence on anthocyanin compositions in sweetpotato leaves. *HortScience*, *40*(1), 176–180.

Klein, A. T., Yagnik, G. B., Hohenstein, J. D., Ji, Z., Zi, J., Reichert, M. D., MacIntosh, G. C., Yang, B., Peters, R. J., Vela, J., Lee, Y. J. (2015). Investigation of the Chemical Interface in the Soybean–Aphid and Rice–Bacteria Interactions Using MALDI-Mass Spectrometry Imaging. *Analytical Chemistry*, *87*(10), 5294–5301. http://doi.org/10.1021/acs.analchem.5b00459

Kompauer, M., Heiles, S., & Spengler, B. (2017). Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4-mum lateral resolution. *Nature Methods*, *14*(1), 90–96. http://doi.org/10.1038/nmeth.4071

Krapp, A., & Stitt, M. (1995). An evaluation of direct and indirect mechanisms for the "sink-regulation" of photosynthesis in spinach: Changes in gas exchange, carbohydrates, metabolites, enzyme activities and steady-state transcript levels after cold-girdling source leaves. *Planta*, *195*(3), 313–323. Retrieved from http://www.jstor.org/stable/23383277

Lebowitz, R. J. (2011, February 9). The Genetics and Breeding of Coleus. *Plant Breeding Reviews*. http://doi.org/doi:10.1002/9781118061008.ch9

Lee, Y. J., Perdian, D. C., Song, Z., Yeung, E. S., & Nikolau, B. J. (2012). Use of mass spectrometry for imaging metabolites in plants. *Plant Journal*, *70*(1), 81–95. http://doi.org/10.1111/j.1365-313X.2012.04899.x

Lemoine, R., Camera, S. La, Atanassova, R., Dédaldéchamp, F., Allario, T., Pourtau, N., Bonnemain, J.-L., Laloi, M., Coutos-Thévenot, P., Maurousset, L., Faucher, M., Girousse, C., Lemonnier, P., Parrilla, J., & Durand, M. (2013). Source-to-sink transport of sugar and regulation by environmental factors. *Frontiers in Plant Science*, *4*(July), 1–21. http://doi.org/10.3389/fpls.2013.00272

Logan, B. A., Stafstrom, W. C., Walsh, M. J. L., Reblin, J. S., & Gould, K. S. (2015). Examining the photoprotection hypothesis for adaxial foliar anthocyanin accumulation by revisiting comparisons of green- and red-leafed varieties of coleus (Solenostemon scutellarioides). *Photosynthesis Research*, *124*(3), 267-274. http://dx.doi.org/10.1007/s11120-015-0130-0

McCormick, A. J., Cramer, M. D., Watt, D. A. (2006). Sink strength regulates photosynthesis in sugarcane. *New Phytologist*, *171*(4), 759–770. http://doi.org/10.1111/j.1469-8137.2006.01785.x

McCormick, A. J., Cramer, M. D., Watt, D. A. (2008). Changes in photosynthetic rates and gene expression of leaves during a source-sink perturbation in sugarcane. *Annals of Botany*, *101*(1), 89–102. http://doi.org/10.1093/aob/mcm258

Nemes, P., & Vertes, A. (2007). Laser Ablation Electrospray Ionization for Atmospheric Pressure, in Vivo, and Imaging Mass Spectrometry. *Analytical Chemistry*, *79*(21), 8098–8106. http://doi.org/10.1021/ac071181r

Nielsen, T.H., Krapp, A., Röper-Schwarz, U., Stitt, M. (2002). The sugar-mediated regulation of genes encoding the small subunit of Rubisco and the regulatory subunit of ADP glucose pyrophosphorylase is modified by phosphate and nitrogen. *Plant, Cell & Environment*, *21*(5), 443–454. http://doi.org/10.1046/j.1365-3040.1998.00295.x

Nilsson, A., Goodwin, R. J. A., Shariatgorji, M., Vallianatou, T., Webborn, P. J. H., Andrén, P. E. (2015). Mass Spectrometry Imaging in Drug Development. *Analytical Chemistry*, *87*(3), 1437-1455. http://doi.org/10.1021/ac504734s

Nguyen, P., & Cin, V. D. (2009). The role of light on foliage colour development in coleus (Solenostemon scutellarioides (L.) Codd). *Plant Physiology and Biochemistry*, *47*(10), 934–945. http://doi.org/10.1016/j.plaphy.2009.06.006

Nguyen, P., Quesenberry, K., & Clark, D. (2008). Genetics of growth habit and development of new coleus (Solenostemon scutellarioides (L.) Codd) varieties with trailing habit and bright color. *Journal of Heredity*, *99*(6), 573–580. http://doi.org/10.1093/jhered/esn054

Osorio, S., Ruan, Y.-L., & Fernie, A. R. (2014). An update on source-to-sink carbon partitioning in tomato. *Frontiers in Plant Science*, *5*(October), 1–11. http://doi.org/10.3389/fpls.2014.00516

Paul, M. J., & Driscoll, S. P. (1997). Sugar repression of photosynthesis: the role of carbohydrates in signalling nitrogen deficiency through source:sink imbalance. *Plant, Cell and Environment*, *20*(1), 110–116. http://doi.org/10.1046/j.1365-3040.1997.d01-17.x

Paul, M. J., Driscoll, S. P., & Lawlor, D. W. (1992). Sink-Regulation of Photosynthesis in Relation to Temperature in Sunflower and Rape. *Journal of Experimental Botany*, *43*(2), 147–153. Retrieved from http://dx.doi.org/10.1093/jxb/43.2.147

Paul, M. J., & Foyer, C. H. (2001). Sink regulation of photosynthesis. *Journal of Experimental Botany*, *52*(360), 1383–1400. Retrieved from http://dx.doi.org/10.1093/jexbot/52.360.1383

Robichaud, G., Barry, J. A., & Muddiman, D. C. (2014). IR-MALDESI Mass Spectrometry Imaging of Biological Tissue Sections Using Ice as a Matrix. *Journal of The American Society for Mass Spectrometry*, *25*(3), 319–328. http://doi.org/10.1007/s13361-013-0787-6

Roitsch, T. (1999). Source-sink regulation by sugar and stress. *Current Opinion in Plant Biology*, *2*(3), 198–206. https://doi.org/10.1016/S1369-5266(99)80036-3

Schmidt, M. A., Barbazuk, W. B., Sandford, M., May, G., Song, Z., Zhou, W., Nikolau, B.J., Herman, E. M. (2011). Silencing of Soybean Seed Storage Proteins Results in a Rebalanced Protein Composition Preserving Seed Protein Content without Major Collateral Changes in the Metabolome and Transcriptome. *Plant Physiology*, *156*(1), 330 LP-345. https://doi.org/10.1104/pp.111.173807

Shiea, J., Huang, M. Z., Lee, C.Y., Yuan, C.H., Beech, I., Sunner, J. (2005). Electrospray-assisted laser desorption/ionization mass spectrometry for direct ambient analysis of solids. Rapid Commun. Mass Spectrom., 19: 3701-3704. https://doi.org/10.1002/rcm.2243

Shroff, R., Schramm, K., Jeschke, V., Nemes, P., Vertes, A., Gershenzon, J., & Svatoš, A. (2015). Quantification of plant surface metabolites by matrix-assisted laser desorption-ionization mass spectrometry imaging: Glucosinolates on Arabidopsis thaliana leaves. *Plant Journal*, *81*(6), 961–972. http://doi.org/10.1111/tpj.12760

Smith, C. A., O'Maille, G., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., Custodio, D. E., Abagyan, R., Siuzdak, G. (2005). METLIN: a metabolite mass spectral database. *Ther Drug Monit*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16404815

Stein, S. E. (1999). An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. *Journal of the American Society for Mass Spectrometry*, *10*(8), 770–781. http://doi.org/10.1016/S1044-0305(99)00047-1

Steyn, W. J., Wand, S. J. E., Holcroft, D. M., & Jacobs, G. (2002). in vegetative tissues : Anthocyanins in unified function a proposed photoprotection. *New Phytologist*, *155*(3), 349–361. http://doi.org/10.1046/j.1469-8137.2002.00482.x

Turgeon, R. (1989). The Sink-Source Transition in Leaves. *Annual Review of Plant Physiology and Plant Molecular Biology*, *40*(1), 119–138. http://doi.org/10.1146/annurev.pp.40.060189.001003

Wardlaw, I. F. (2006). Tansley Review No. 27 The control of carbon partitioning in plants. *New Phytologist*, *116*(3), 341–381. http://doi.org/10.1111/j.1469-8137.1990.tb00524.x

Wu, X., Gu, L., Prior, R. L., & McKay, S. (2004). Characterization of Anthocyanins and Proanthocyanidins in Some Cultivars of Ribes, Aronia, and Sambucus and Their Antioxidant Capacity. *Journal of Agricultural and Food Chemistry*, *52*(26), 7846–7856. http://doi.org/10.1021/jf0486850

Zavalin, A., Yang, J., Hayden, K., Vestal, M., & Caprioli, R. M. (2015). Tissue protein imaging at 1 μm laser spot diameter for high spatial resolution and high imaging speed using transmission geometry MALDI TOF MS. *Analytical and Bioanalytical Chemistry*, *407*(8), 2337–2342. http://doi.org/10.1007/s00216-015-8532-6

Table 1. Flavonoids identified by ELDI-MS. Peak assignments were based on matches with entries in the METLIN Database (Smith et al., 2005). Code: * denotes MS/MS peak matches with METLIN Database, $^{+}$ denotes metabolites identified through LCMS.

High Res. *m/z*	Assigned ID	Δppm of ID	Response to Shading	Major Fragment Ions	Signal Intensity
255.0679	Catechine + H^+ -2H_2O	6	None	Background interferences	5e4
271.0627	*Apigenin + H^+	9	None	243, 229, 225, 156, 145, 119	1e6
287.0585	*Cyanidin (M^+)	12	None	269, 241, 213, 185, 157, 137	7e5
291.0899	Catechine + H^+	12	None	Too low abun.	2e4
399.0434	*Flavone + K^+	10	None	381, 371, 353, 337, 287, 219, 201, 175	1e6
441.0725	Cyanidin Flavonoid + (M^+) +H^+	3	None days 7-14, decreased days 21-28	287	2e5
447.0918	*Apigenin Glucuronide + H^+	0.4	None	271, 163	7e5
458.0823	Flavone + 2H^{2+}	4	None	399, 371, 353, 219, 163, 147	8e4
476.0907	Cyanidin-coumaroyl glucoside-dimalonylglucoside (M^+) + Na^+	6	Decreased	287	7e4
483.0781	Cyanidin Flavonoid (M^+)		None	287, 203, 163, 147	8e5
489.1057	Apigenin Flavonoid + H^+ OR + H^+ - H_2O	4	None days 7-21, decreased day 28	471, 399, 271, 163	5e5
505.1016	Flavone Glucoside + H^+– H_2O	6	None		5e4
511.0744	Cyanidin Flavonoid (M^+)		Decreased	437, 353, 313, 287, 271, 179	7e4
527.0515	Cyanidin Flavonoid (M^+)	3	Decreased	369, 309, 287, 163	1e5
533.2350	Flavonoid		None	287, 271, 219, 201, 163, 147	2e5
535.1069	*Cyanidin malonylglucoside (M^+)	2	None days 7-21, decreased day 28	287, 163	6e4
549.2095	Cyanidin Flavonoid (M^+)		None	369, 331, 287, 271, 219, 201, 163, 147	2e5
557.0940	Flavonoid + Na^+	6	None	Interference from lock mass	8e4
593.1339	Flavonoid + H^+– H_2O	7	None	Too low abun.	2e4
595.1447	*Cyanidin coumaroylglucoside (M^+)	0.1	Decreased	549, 287, 163, 147	3e5
619.2228	Apigenin Flavonoid + H^+		Increased	575, 533, 271, 163	2e5
635.2041	Apigenin Flavonoid + H^+		Increased	591, 549, 271, 163	4e5
697.1607	Flavonoid + H^+	0.4	Decreased	Too low abun.	1e4
743.1454	Apigenin Flavonoid + H^+		None	271, 163, 147	4e4
757.1965	*+Cyanidin-coumaroylglucoside-glucoside (M^+)	1	None	595, 449, 399, 287, 271, 163, 147	8e5

Table 1. (cont'd)

773.1890	Flavonoid + H^+	4	None	Too low abun.	1e4
787.2041	Flavonoid + H^+	4	None	Too low abun.	2e4
799.2031	Cyanidin Flavonoid $M^+ – H_2O$ or $M^+ – 2H_2O$	0.1 - 6	Decreased	621, 595, 535, 527, 441, 287, 163, 147	7e4
813.1613	Cyanidin Flavonoid (M^+)		None	795, 287, 163	7e4
827.1994	Flavonoid + H^+	4-5	Decreased	Too low abun.	2e4
829.1490	Cyanidin Flavonoid (M^+)		None	595, 287, 271, 163, 147	5e4
843.1970	$^{*+}$Cyanidin-coumaroyl glucoside-malonylglucoside (M^+)	0.4	Decreased	799, 595, 535, 287, 163, 147	4e6
859.1876	$^{*+}$Cyanidin- caffeoylglucoside-malonylglucoside (M^+)	5	Decreased	843, 489, 287, 271, 163	7e4
865.1617	Flavonoid + $H^+ – H_2O$	0.1-7	Decreased	839, 821, 677, 635, 617, 575, 531, 513	1e5
873.2090	Cyanidin Flavonoid (M^+)	0.6+	None	843, 693, 595, 535, 489, 287, 163, 147	5e4
883.1759	Dicatechine Flavonoid + H^+	4	None	865, 839, 821, 677, 635, 617, 575, 531, 513	2e5
885.2034	Flavonoid		Decreased	867, 595, 577, 287, 271, 163, 147	9e4
899.1660	Dicatechine Flavonoid + H^+	0.5	None	881, 855, 837, 813, 693, 651, 633, 591, 547	3e5
929.1966	$^{*+}$Cyanidin-coumaroyl glucoside-dimalonylglucoside (M^+)	1	Decreased	843, 621, 595, 489, 287, 163, 147	1e6
951.1619	Cyanidin Flavonoid (M^+)		Decreased	865, 843, 741, 635, 549, 489, 447, 399, 287	5e4

Table 2. Carbohydrates identified by ELDI-MS. Peak assignments were based on matches with entries in the METLIN Database (Smith et al., 2005). Code: * denotes MS/MS peak matches with METLIN Database, # denotes metabolites identified through GCMS.

High Res. *m/z*	Assigned ID	Δppm of ID	Response to Shading	Major Fragment Ions	Signal Intensity
159.0021	Tetrose + K^+	20	None	Localized Signal	8e4
189.0158	#Pentose + K^+	0.1	None	Localized Signal	3e4
197.0433	*Organic Acid Glucoside + H^+ + Na^+	0.01	None	161, 145, 119, 101	1e6
203.0560	*#Hexose (glucose) + Na^+	16	No abundance day 7, increased days 14-28	185, 167, 157, 137, 123, 111	1e3 - 5e5
219.0310	*#Hexose (glucose) + K^+	20	No abundance day 7, increased days 14-28	201, 183, 151, 123, 111	5e3 - 6e5
233.0671	*Heptose + Na^+	16	None	215, 205, 179	5e5
249.0414	*Heptose + K^+	17	None	231, 213, 195, 125	4e5
255.1064	Hexose-glycerol + H^+	4	None	Too low abun.	1e4
277.0901	Hexose-glycerol + Na^+	0.1	None	Background interferences	5e4
312.9720	Hexose-Phosphate + K^+	0.01	Low abundance day 7, increased days 14-28		3e3 – 9e4
337.0585	*Caffeoylglucarate + H^+ - $2H_2O$	5	None	319, 185, 163	7e4
349.0923	*Organic Acid Glucoside + Na^+	8	Increased	331, 313, 267, 163, 149	9e4
357.0586	Hexose-glycerol Phosphate + Na^+	8	None, decreased in abundance day 28		4e3 – 1e5
365.0647	*Organic Acid Glucoside + K^+	3	Increased	347, 333, 329, 163	1e5
371.0475	*Galloylglucose + K^+	26	None	353, 327, 237, 219	1e5
381.0786	*#Sucrose + K^+	2	None	363, 345, 335, 219, 163, 147	7e4

Table 3. Amino acids and organic acids identified by ELDI-MS. Peak assignments were based on matches with entries in the METLIN Database (Smith et al., 2005). Code: * denotes MS/MS peak matches with METLIN Database, # denotes metabolites identified through GCMS.

High Res. *m/z*	Assigned ID	Δppm of ID	Response to Shading	Major Fragment Ions	Signal Intensity
72.0454	Alanine + H^+ - H_2O	6	None	Too low *m/z*	7e3
73.0309	#Lactic Acid + H^+ – H_2O	26	None	Too low *m/z*	1e4
81.0707	Caproic Acid + H^+ – $2H_2O$	3	None	Too low *m/z*	1e4
83.0145	#Succinic Acid + H^+ - $2H_2O$	7	None	Too low *m/z*	8e3
86.0588	GABA + H^+ - H_2O	20	None	Too low *m/z*	3e4
98.0198	#Glycine + Na^+	14	None	Too low abun.	6e3
98.0585	Proline H^+ - H_2O	21	None	Too low abun.	9e3
106.0494	*#Serine + H^+	4	None	88	1e4
114.0907	#Leucine + H^+ – H_2O	10	None	Background interferences	1e3 – 2e5
115.0029	#Tartaric Acid + H^+ – $2H_2O$	6	None	Localized Signal	2e4
116.0365	*Aspartic Acid + H^+ – H_2O	14	Increased	88	3e4
118.0863	*Valine + H^+	~0	None	100, 72	3e4
120.0658	*Threonine + H^+	2	None	102	6e4
129.0346	*#p-Coumaric Acid + H^+-$2H_2O$	~0	None	119, 91	2e4
130.0520	*#Pyroglutamic Acid + H^+	16	None	84	4e5
145.0322	*#Caffeic Acid + H^+ - $2H_2O$	18	None	135, 117, 107, 89, 79	1e5
147.0485	*#p-Coumaric Acid + H^+-H_2O	26	None	119, 91	4e5
147.0777	*Glutamine + H^+	8	None days 7-21, Increased day 28	130, 84	2e4
155.0423	*Asparagine + Na^+	2	None	109	2e4
156.0809	*Histidine + H^+	26	None days 7-14, Increased days 21-28	110, 93, 83	5e4
157.0133	#Citric Acid + H^+ – $2H_2O$	6	None	Too low abun.	9e3
163.0430	*#Caffeic Acid + H^+-H_2O	20	None	145, 135, 117, 107, 89, 79	8e5
170.0563	*#Leucine + K^+	8	None	124	5e4
175.1200	*Arginine + H^+	5	Increased	158, 130, 116, 112, 70	1e5
219.0019	*#Caffeic Acid + K^+	16	None	201, 183, 173, 161, 129	3e5
230.9883	*#Citric Acid + K^+	6	None	213, 195	2e4

CHAPTER 5. MAIZE *GLOSSY2 AND GLOSSY2-LIKE* GENES; ARE THEY FUNCTIONAL HOMOLOGS OF ARABIDOPSIS *ECERIFERUM2*?

Manuscript in preparation to be submitted to Plant Physiology

[1]Liza Esther Alexander[a,b], Aeriel Davis[c], Michael A. Schelling[a], Xiaobin Zheng[a,b], Ludmila Rizhsky[a,b], Yozo Okazaki[d], Kazuki Saito[d], Basil J. Nikolau[a,b].

Abstract

Plant epidermal cells express unique molecular machinery that juxtaposes the assembly of intracellular lipid components and the unique extracellular cuticular lipids that are unidirectionally secreted to plant surfaces. In maize (*Zea mays*), mutations at the *glossy2 (gl2)* locus affects the deposition of extracellular cuticular lipids. Sequence-based genome scanning identified a novel homolog in the maize genome, we term *Gl2-like*. Sequence homology identifies that both the *Gl2-like* and *Gl2* genes are members of the BAHD superfamily of acyltransferases, with close sequence homology to the Arabidopsis *CER2* gene. Transgenic experiments demonstrate that *Gl2-like* and *Gl2* functionally complement Arabidopsis *cer2* mutation, with differential impacts on the epicuticular lipids and on the lipidome of the plant. *Gl2* and *Gl2-like* have different effects at the longer alkyl chain lipids, specifically at the 32-carbon chain length. Site-directed mutagenesis of the putative BAHD catalytic HXXXDX motif indicate that *Glossy2-like* requires the BAHD-type

[a] Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA, USA.
[b] Center for Metabolic Biology, Iowa State University, Ames, IA, USA.
[c] Departments of Genetics, Development, and Cell Biology and Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA, USA.
[d] Metabolomics Research Group, Center for Sustainable Resource Science, RIKEN, Yokohama, Kanagawa, Japan.

catalytic capability to fully complement the *cer2* function, but *Glossy2* can accomplish this without the need for this catalytic motif.

Introduction

Extracellular epicuticular lipids are constituents of a protective hydrophobic structure, the cuticle, which covers aerial organs of all terrestrial plants. The cuticle plays important roles in many plant-environment interactions, including controlling water status (Kerstiens, 1996; Riederer, 2006), responses to abiotic stresses (Long, 2003), interactions with biotic pathogens (Kolattukudy, 1985; Jenks et al., 1994), and in defining organ boundaries during development (Yephremov et al., 1999; Sieber et al., 2000; Kurdyukov et al., 2006). The extracellular cuticular lipids are chemically and physically arranged in distinct layers (Kolattukudy, 1965; Jellings and Leech, 1982; Jetter et al., 2000), with the epicuticular lipids being primarily linear alkyl chains, including very-long chain fatty acids (VLCFAs), hydrocarbons, ketones, primary and secondary alcohols, aldehydes and wax esters; in addition, non-alkyl, terpene-type specialized metabolites are also components of the epicuticle (Martin and Juniper, 1970; Kolattukudy, 1976; Tulloch 1976). These lipids are deposited on and embedded within a lipophilic cutin polymer matrix (a polymer of esterified ω–hydroxy and epoxy-C16 and C18 fatty acids, glycerol, and α,ω–dicarboxylic acids) (Kolattukudy, 2001; Heredia, 2003; Bonaventure et al., 2004; Franke et al., 2005; Pollard et al., 2008).

The molecular aspects of epicuticular lipid biogenesis have been greatly facilitated by forward genetic approaches that use e*ceriferum (cer)* mutants of Arabidopsis (Koornneef et al., 1989), *glossy (gl)* mutants of maize (Schnable et al., 1994) and tomato (Vogg et al., 2004; Leide et al., 2007) and *wax crystal-sparse leaf (wsl)* mutants of rice (Yu et al., 2008; Wang et al., 2017). These studies have been primarily interpreted in the context of a metabolic model that had been

proposed from earlier physiological/biochemical studies (Bianchi et al., 1985; Post-Beittenmiller, 1996). At the core of this metabolic model is the process of fatty acid elongation, which feeds two fatty-acid modification pathways: a reductive pathway that generates fatty aldehydes, primary alcohols, and wax esters; and a decarboxylative pathway that converts the common aldehyde intermediate to hydrocarbons, and ultimately ketones, and secondary alcohols.

In maize both fatty acid modification pathways are differentially expressed among different organs. In leaves of maize seedlings specifically, the reductive pathway predominates, and the expression of this metabolic network is under the control of the juvenile-to-adult phase transition, mediated by the transcription factor *Gl15* (Moose and Sisco, 1996). The *glossy* mutations that have been characterized since the early-1900s have been identified via phenotypic screens of seedling leaves, which therefore primarily affect the fatty acid elongation-reductive pathway. More recently, studies on epicuticular lipid composition of maize silks demonstrate that up to ~90% of the lipids covering this tissue are composed of hydrocarbons, implying that the fatty acid elongation-decarboxylative pathway is active in this tissue (Perera et al., 2010; Loneman et al., 2017).

Glossy2 is exemplary of such a product of forward genetics, initially identified in 1928 (Hayes and Brewbaker, 1928) as a mutant that "beads" water droplets on seedling leaves, a common phenotype of all *glossy* mutants. The initial characterization of the epicuticular lipid profiles of *gl2* (Bianchi, 1975) identified deficiencies in epicuticular lipids that are derived from the fatty acid elongation-reductive pathway, specifically a block in the fatty acid elongation process between 30 and 32 carbon chain length. The homology of *glossy2* to the Arabidopsis *CER2* became apparent when both loci were molecularly cloned (Tacke et al., 1995; Xia et al., 1996); the two genes encode proteins that share ~60% sequence similarity. The *cer2* mutation

affects the bright-green appearance of Arabidopsis stems (Koornneef et al., 1989), with an underlying reduction in the fatty acid elongation-decarboxylative pathway, with apparent block in the elongation of C26 or C28 fatty acids to C30 fatty acid (Mcnevin et al., 1991; Jenks et al., 1995).

At the time of initial identification, these two proteins (GL2 and CER2) defined novel sequences (Tacke et al., 1995; Xia et al., 1996), which ultimately became archetypal of the BAHD class of enzymes (St-Pierre and Luca, 2000; D'Auria, 2006). BAHD enzymes are acyl-CoA-dependent acyltransferases that catalyze the acylation of alcohols or amine groups, forming ester and amide groups in the assembly of a large number of specialized metabolites (St-Pierre and Luca, 2000; D'Auria, 2006). These enzymes have been phylogenetically categorized in to five clades (Clades I-V), with GL2 and CER2 being members of Clade II (D'Auria, 2006). The specific biochemical and physiological functions of the GLOSSY2/CER2 Clade II BAHD enzymes remains uncharacterized. More recently, yeast reconstitution experiments suggest that *CER2* (Haslam et al., 2012) and its rice homolog, *OsCER2* (Wang et al., 2017) associates with the fatty acid elongase (FAE) system, but the role of BAHD catalytic capability for this association is unclear.

In this study, we identified a novel *Glossy2-like* gene, which is an paralog of *Gl2,* and explored whether these two products of an evolutionary gene duplication event have provided a template for potential neofunctionalization (Renny-Byfield and Wendel, 2014). Specifically, we engaged a transgenic strategy and evaluated and compared the functional ability of *Glossy2* and *Glossy2-like* to complement the Arabidopsis *cer2* mutant. The resulting transgenic lines were used to evaluate the effect of these genetic manipulations on the epicuticular lipid profiles and also on the broader lipid metabolic network (i.e., cutin and lipidome). These experiments demonstrate that both maize *Glossy2* and *Glossy2-like* genes can complement the Arabidopsis *cer2* mutant,

however, the two genes have functionally drifted, with *Gl2* acquiring additional capabilities as compared to *Gl2-like* and *CER2*.

Results

Computational identification and characterization of the *Glossy2-like* gene

The phylogenetic diversity of the *Glossy2* (*Gl2,* GRMZM2G098239*)* gene was explored by sequence-based homology searches of both the Arabidopsis and maize genomes (Altschul et al., 1997; Alonso et al., 2003; Andorf et al., 2016). This revealed that GL2 shares high sequence similarity with an uncharacterized maize gene, we term *Glossy2-like* (*Gl2-like,* GRMZM2G315767), and similarities with three Arabidopsis homologs, *CER2* (At4g24510), *CER2-LIKE1* (At4g13840), and *CER2-LIKE2* (At3g23840) (Figure 1) (Xia et al., 1996; Pascal et al., 2013). The GL2 and GL2-LIKE proteins share 50% sequence identity and 63% sequence similarity with each other, and ~30% sequence identity and 50% sequence similarities with the CER2-family of proteins. This sequence similarity includes the conserved acyl-CoA-dependent acyl transferase domain, HXXXDX, which is typical of the BAHD enzyme family (D'Auria, 2006).

Transgenic expression of GLOSSY2 and GLOSSY2-LIKE modifies the *eceriferum* phenotype of the Arabidopsis *cer2* mutant stems

The functionality of the *Gl2* and *Gl2-like* genes was evaluated by their ability to genetically complement the *cer2* mutant of Arabidopsis. ORFs encoding these two proteins were transgenically expressed under the transcriptional control of the constitutive 35S promoter in wild-type Arabidopsis and Arabidopsis mutant lines homozygous for the *cer2-5* allele. Expression of the *Gl2* and *Gl2-like* transgenes was confirmed at the protein and mRNA levels via western blot analysis (Figure 2A) of extracts using a GLOSSY2 antibody, and by RT-PCR analysis of RNA

isolated from these plants (Figure 2B), respectively. As expected, the GL2 antibody detected a 46-kDa-polypeptide band in extracts from maize silk tissues, and this protein band was also detected in extracts prepared from Arabidopsis transgenic lines that are expressing the *Gl2* gene (i.e., genotype: *ZmGl2* in WT and *ZmGl2* in *cer2-5*), but this protein band is absent from the control non-transgenic Arabidopsis plants.

Using RT-PCR, with RNA-template preparations, the *Gl2-like* transcript was detected in the appropriate Arabidopsis transgenic lines (i.e., *ZmGl2-like* in WT, *ZmGl2-like* in *cer2-5*), and this transcript was undetectable in wild-type, and *cer2-5* mutant control lines (Figure 2B). In both sets of these transgenic experiments, three replicate lines from each of three independent transformation events were propagated and subjected to analyses, and the resulting data were compared to the appropriate wild-type controls not expressing the transgenes.

Typical of *cer2* mutants, the stems of the *cer2-5* mutant show the *eceriferum* phenotype, presenting bright green stems (Koornneef et al., 1989), rather than the dull green appearance of the wild-type plants (Figure 2C and 2D). This phenotype is associated with epicuticle-deficiency and is indicative of changes in the epicuticular surface lipid composition (Koornneef et al., 1989). As with the transgenic expression of the CER2 protein, the transgenic expression of the GL2-LIKE protein in the *cer2-5* mutant background, fully restores the stem *eceriferum* phenotype to the dull green, wild-type appearance (Figure 2D).

In contrast, the transgenic expression of the GL2 protein in the *cer2-5* mutant only partially restored this phenotype to the wild-type appearance (Figure 2C). As control experiments we transgenically expressed the *Gl2* or *Gl2-like* transgenes in the wild-type background, and this did not alter the phenotype of the stems (i.e., they retained the dull green, wild-type appearance).

As is expected from prior studies (Xia et al., 1996), scanning electron microscopy (SEM) examination of the stems identify the crystalloid structures of the cuticle, and these are devoid from the surfaces of Arabidopsis *cer2*-5 mutant stems, and the overexpression of *CER2* in this mutant background induces their reappearance and they have similar appearance to those of wild-type plants (Figure 3). The parallel SEM examination of the stems of the plants expressing the *Gl2-like* transgene in either wild-type or *cer2-5* mutant background are similar to those of wild-type plants (Figure 3). In contrast, the crystalloids on those plants expressing the *Gl2* transgene, in either the wild-type or *cer2-5* mutant background, occur at a lower density, and the structures of these crystalloids are more irregular, flattened, and have a thinner flake-like appearance than the wild-type control (Figure 3).

Transgenic expression of *Glossy2* and *Glossy2-like* genes rescues the epicuticular lipid-deficiency chemotype of the *cer2* mutant

Extracellular epicuticular lipids were extracted from stems of the different genotypes developed in this study and subjected to GC-MS analysis, and the resulting data are presented in Figure 4, Supplemental Figures 1 and 2, and Supplemental Table 1. The extractable epicuticular lipid load on stems of the *cer2-5* mutant is about half of the wild-type stems (Supplemental Figure 2). The transgenic overexpression of either the *Gl2* or the *Gl2-like* transgenes in the *cer2-5* mutant increased the lipid load by about two-fold to near wild-type levels. The transgenic expression of *Gl2-like* in the wild-type background also increased the total epicuticular lipid load by ~40% of the wild-type load, whereas the expression of the *Gl2* transgene had no such effect on the total epicuticular lipid load of the wild-type stems.

Figure 4 shows the compositional changes in the epicuticular lipid profiles of these stems; these data are focused on alkyl chain lengths of 26-carbons and greater, and accumulation of the

short chain length alkyl chains were negligible and were only subtly affected by these transgenic events. The *cer2-5* mutation primarily affects the accumulation of the major components of the epicuticular lipids, which are nearly eliminated in the mutant; these constituents are alkyl derivatives of C30 fatty acid, namely, C29 alkane, C29 secondary alcohol, and C29 ketone. These decreases in accumulation are associated with a partial compensatory increase in shorter chain fatty acid derivatives, predominantly the C26 primary alcohol. The transgenic expression of either *Gl2* or *Gl2-like* transgenes in the *cer2-5* mutant background results in an epicuticular lipid compositional profile that is near identical to the wild-type profile. This is specifically associated with increased accumulation of the C29 alkane, C29 secondary alcohol, C29 ketone, and decreased accumulation of the C26 primary alcohol (Figure 4A).

In addition to these genetic compensatory effects on the epicuticular lipid profiles, the expression of these two maize transgenes induced novel changes to the profiles that are not present in the wild-type plants. In particular, *Gl2* expression in *cer2-5* mutant induces the formation of longer alkyl chains; namely derivatives of the C32 fatty acid, which includes C31 alkane, C31 secondary alcohol, and C31 ketone. The appearance of these novel metabolites is also observed when the *Gl2* transgene is over-expressed in wild-type background (Figure 4).

These novel components are not manifest by the transgenic expression of either the *CER2* or the *Gl2-like* genes in either the wild-type or the *cer2-5* mutant (Figure 4). Collectively, these biochemical changes establish that both the GL2 and GL2-LIKE proteins are functional homologs of CER2 and fully complement the biochemical deficiency associated with the *cer2* mutation, however the *Gl2* transgene expresses additional capabilities by enabling the production of even longer chains constituents than normal (up to 32-carbon fatty acids, and their alkyl derivatives).

Mutations in the putative BAHD catalytic domain demonstrate differences in the *in vivo* functionality between *Glossy2 and Glossy2-like* genes

The maize GL2 (Tacke et al., 1995), and the homologous Arabidopsis CER2 (Negruk et al., 1996; Xia et al., 1996) proteins ultimately became archetypal of the BAHD family of enzymes that catalyze acyltransferase reactions (D'Auria, 2006). Two common structural features are important to the function of BAHD enzymes, which have been identified through structural analysis and site-directed mutagenesis studies (Ma et al., 2005; D'Auria, 2006; Unno et al., 2007; Garvey et al., 2008). One of these is the HXXXDX-motif, which contains the catalytic His-residue that is responsible for deprotonating the alcohol or amine acyl acceptor-substrate, and is thus crucial to catalysis. The second conserved sequence is the DFGWG-motif, which appears to be structurally important for substrate selectivity (Suzuki et al., 2003; Bayer et al., 2004; Ma et al., 2005; Unno et al., 2007; Garvey et al., 2008). The location and identity of the HXXXDX-motif in the primary sequences of GL2 and GL2-LIKE proteins is easily recognized whereas, the identity of the DFGWG-motif is not apparent (Figure 1).

Additional insights on the interaction between the potential substrates and the BAHD catalytic domain was obtained with the computational predicted structural model of GL2 and GL2-LIKE, generated by *Phyre2* (protein homology/analogy recognition engine, (Kelley et al., 2015). Using these structural models, the *3DLigandSite* algorithm (Kelley and Sternberg, 2009) identified the *Sorghum* hydroxycinnamoyl transferase (HCT) as the best BAHD structural template for GL2 and GL2-LIKE proteins. Using the experimentally determined structure of the substrate-enzyme complex of HCT, we identified that the last "X" residue of the HXXXDX-motif of GL2 and GL2-LIKE (i.e., Ile-174 and Ile-190, respectively) may also interact with the potential substrate.

The role of the HXXXDX-domain in GL2 and GL2-LIKE functions were tested by site-directed mutagenesis. Specifically, we generated three point-mutants of the GL2 and GL2-LIKE proteins, by substituting Ala for each of the critical residues in the HXXXDX motif (i.e., GL2(H169A), GL2(D173A), GL2(I174A), and GL2-LIKE(H185A), GL2-LIKE(D189A) and GL2-LIKE(I190A)). Each of these mutants were transgenically expressed in the *cer2-5* mutant and wild-type Arabidopsis plants and their stem cuticular lipid profiles were analyzed to assess the ability of the mutant GL2 and mutant GL2-LIKE proteins to complement the *cer2-5* mutation (Figure 5, and Supplemental Table 2).

All three GL2 point mutants, did not affect the ability of these transgenes to complement the *cer2-5* chemotype; namely either the wild-type GL2 transgene or any one of the three GL2 point mutants that should have destroyed the BAHD catalytic capability are still capable of reversing the *cer2-5* reduction in epicuticular lipid accumulation (Figure 5A). In contrast, whereas the wild-type *Gl2-like* transgene can also fully complement the *cer2-5* chemotype, and reverse the epicuticular lipid load on these stems, the H185A, and I190A point mutants could not fully support this genetic complementation (Figure 5B). These results demonstrate that the HXXXDX BAHD catalytic motif is not required for the ability of the *Glossy2* gene to functionally replace the *CER2* function, however this motif is required for the ability of the *Glossy2-like* gene to fully replace the *CER2* function.

Lipidomics and cutin profiling reveal alterations in VLCFAs and hydroxy-VLCFAs

Characterizations of the phenotypic and chemotypic effects of the *eceriferum* mutations have primarily focused on the epicuticular lipids. However, the primary genetic lesion that determines these extracellular lipid surface traits occurs in the intracellular lipid metabolic processes that underlie the origins of the epicuticular lipid material. These processes are primarily

associated with the intracellular membranes (ER and possibly plasma membrane) that house the biochemical reactions associated with VLCFA biosynthesis and down-stream reactions (Bernard and Joubès, 2013). Furthermore, in addition to epicuticular lipid biosynthesis, VLCFA metabolism is also juxtaposed with other lipid metabolic pathways, including phospholipids, sphingolipids, storage lipids, and suberin (Li-Beisson et al., 2010). Additionally, the epicuticular lipids are closely associated and integrated with the polymeric cutin matrix, which is assembled by covalent ester bonds, which are the types of bonds that BAHD enzymes produce (D'Auria, 2006). Therefore, we profiled and compared the intracellular lipidomic pools and of the monomers of the cutin matrix in the stems of the different genotypes generated in this study.

The *cer2* mutation does not have major effects on the composition of cutin monomers. However, these preparations contain VLCFA components (i.e., 2-hydroxy-VLCFAs), which are probably associated with sphingolipids that co-purify with the cutin preparations (Molina et al., 2006). Regardless of their complex-lipid origins, because they are FAE-generated products we evaluated the effect of *Gl2* and *Gl2-like* transgenic expression on their abundance. Specifically, in the *cer2-5* mutant there are reductions in accumulation of these C22- and C24- fatty acids and their 2-hydroxylated derivatives, and an increase in the C26 fatty acid and the corresponding 2-hydroxylated derivative (i.e., metabolite #163, #165, #167, #176, #178, and #181, Supplemental Figure 3A). Transgenic expression of *Gl2* in the *cer2-5* mutant background reversed these alterations in the abundance of these VLCFA components, and in some cases further enhanced their accumulation to double or quadruple (i.e., metabolites #163, #165, #176, #178, and #179, Supplemental Figure 3C) their accumulation as compared to the levels found in the wild-type controls. In contrast, to the effect induced by the transgenic expression of *Gl2*, the transgenic expression of *Gl2-like* caused very minor alterations in the levels of cutin monomers and the

VLCFA components associated with the cutin preparations, irrespective of whether *Gl2-like* is expressed in a wild-type or *cer2-5* mutant background (Supplemental Figure 3E and 3F).

Additional insights of the effect of transgenic expression of *Gl2* and *Gl2-like* on lipid metabolism was provided by lipidomics analyses of Arabidopsis stems. Lipid extracts were analyzed using LC-qTOF, which enabled the identification and relative quantification of 1372 analytes, 153 of which were chemically identified. The latter includes phosphoglycerolipids (29), glycolipids (35), isoprenoids (24), acylated glycolipids (8), sterol lipids (12), storage lipids (33) and free fatty acids (12) (Supplemental Table 3). The most striking of these alterations are the change to the accumulation of the free fatty acids (i.e., metabolite #148 to #153, Supplemental Figure 3A-F). The *cer2-5* mutation blocked the formation of C30 fatty acid (metabolite #150), resulting in the higher accumulation of the precursors C26 and C28 fatty acids (i.e., metabolite #148 and #149, Supplemental Figure 3A). The transgenic expression of maize *Gl2* and *Gl2-like* in the *cer2-5* mutant reversed these effects (Supplemental Figure 3C and 3E). Moreover, *Gl2* has additional capabilities, inducing the increased accumulation of C32 fatty acid (i.e., metabolite #151, Supplemental Figure 3C and 3D). This latter effect correlates with the increased levels of C31 alkyl derivatives that were detected in the epicuticular lipid profiles of wild-type and *cer2-5* mutant stems overexpressing *Gl2* transgene (Figure 4).

Another lipidomic alteration that is consistent with the shared functionality between *cer2*, *Gl2* and *Gl2-like* is the alteration in the accumulation of sterol esters; i.e., β-sitosterol esterified with linoleic or linolenic acid (i.e., metabolite #106 and #107, Supplemental Figure 3A, 3C and 3E). The accumulation of these two sterol esters is increased in the *cer2-5* mutant, as compared to the wild-type, and it is decreased back to normal levels when *Gl2* was transgenically expressed in the mutant; in contrast, the transgenic expression of *Gl2-like* does not have this effect

(Supplemental Figure 3C and 3E). An additional similar genetic modulation is the accumulation pattern of the glycosylceramides that utilize a 2-hydroxylated C26 fatty acid building block (i.e., metabolite #58 and #59, Supplemental Figure 3A). The accumulation of these metabolites is increased in the *cer2-5* mutant, and this effect is reversed by the transgenic expression of both *Gl2* and *Gl2-like* (Supplemental Figure 3C and 3E).

In addition, several of the chemically undefined 1,219 analytes (Supplemental Table 3) show an accumulation pattern that could be revealing of *Gl2* or *Gl2-like* function, particularly if their chemical identities are determined. In this discussion these analytes are identified via the *m/z* values of the individual ions detected by LC-qTOF in either negative or positive ion modes. These ions are classified into 5 broad categories: Category A are ions whose accumulation is decreased by the *cer2-5* mutation, but is compensated by the expression of either *Gl2* or *Gl2-like* or *CER2* (negative ion-mode: 33 metabolites; positive ion-mode: 19 metabolites); Category B are ions whose accumulation is increased by the *cer2-5* mutation, and is compensated by the expression of either *Gl2* or *Gl2-like* or *CER2* (negative ion-mode: 34 metabolites; positive ion-mode: 28 metabolites); Category C are ions whose accumulation is increased or decreased by the *cer2-5* mutation, and expression of either *Gl2* or G*l2-like* results in the opposite over-compensation in accumulation (negative ion-mode: 10 metabolites; positive ion-mode: 1 metabolite); Category D are ions whose accumulation is unaltered by the *cer2-5* mutation, but expression of either *Gl2* or G*l2-like* results in altered accumulation pattern (negative ion-mode: 218 metabolites; positive ion-mode: 58 metabolites); and Category E are ions whose accumulation patterns are differentially affected between *Gl2* and *Gl2-like* expression (negative ion-mode: 56 metabolites; positive ion-mode: 6 metabolites). Ions that fall in the first four categories (A-D) are consequences of functions that overlap between *CER2* and either *Gl2* and *Gl2-like*, with the former two categories (A and B)

indicating exact overlap, whereas the latter two categories (C and D) may be an indicator of a quantitative difference in that common function. Ions that fall into the 5th category (E) would be indicators of differences in function between either *Gl2* or *Gl2-like*, a product of neofunctionalization following the genes duplication that gave rise to the 2 homologs.

These effects on the stem lipidomes associated with the transgenic addition of the *Gl2* function into a heterologous host (Arabidopsis), was compared to the effect of removing this function in the homologous host, i.e., maize *gl2* mutant seedling leaves. The *gl2* mutation primarily affects the accumulation of the major components of the epicuticular lipids on seedling leaves, reducing the levels of C32 primary alcohol and C32 aldehyde. These decreases in accumulation are associated with only a partial compensatory increase in C28 primary alcohol and aldehyde (Supplemental Figure 4). Parallel lipidomics analyses indicate that similar to the *cer2-5* mutation in Arabidopsis, mutation at the *gl2* locus, primarily affects VLCFA accumulation of the intracellular lipidome of maize seedlings, specifically decreasing the level of C32 fatty acid, coupled with the higher accumulation of the C26 and C28 fatty acid precursors (metabolite #164 and #165, Supplemental Figure 5). Another minor effect that mirrors the Arabidopsis transgenic results is the alteration in the levels of glycosylceramides that utilize hydroxylated C26 fatty acid building block (metabolite #60, Supplemental Figure 5). Analogous comparison could not be conducted with *gl2-like* gene due to the lack of a mutant allele at this maize locus.

Discussion

The extracellular cuticular lipids that coat the outer surfaces of the aerial organs of terrestrial plants (Fernandes et al., 1964; Kolattukudy 1980) are specifically produced by the epidermal cell layer of these organs. Because epidermal cells account for only about 10% of the cellular population of these aerial organs (Jellings and Leech, 1982), elucidating the molecular and

biochemical mechanisms that regulate their biogenesis is confounded by the other 90% of the cell population that is not involved in these processes. This technical barrier has been partially overcome by utilizing forward genetic approaches to identify and characterize genes involved in extracellular cuticular lipid accumulation. This strategy has been successful in the isolation of the causative genes that generate the epicuticular lipid phenotypes, e.g., the molecular identification of the *gl2* or *cer2* genes (Tacke et al., 1995; Xia et al., 1996). Yet in many cases, the exact mechanisms by which these gene products affect epicuticular lipid deposition is still unknown.

We specifically took a transgenic strategy to characterize the maize *Glossy2* gene, and the homologous *Glossy2-like* gene, which shares 63% sequence similarity. These two genes probably arose via a gene duplication mechanism, although the former is known to be involved in epicuticular lipid biosynthesis (Hayes and Brewbaker, 1928; Bianchi, 1975; Tacke et al., 1995), the latter gene may have roles that are distinct from that of *Gl2*. These transgenic experiments utilized Arabidopsis as the vehicle for these evaluations, testing for the ability of each transgene to compensate for a missing function provided by the homologous Arabidopsis *CER2* gene, and in parallel the effect of each transgenic event in the wild-type background.

The ability of *Glossy2* and *Glossy2-like* to contribute to extracellular epicuticular surface lipid deposition

The functionality of *Gl2* in epicuticular lipid biosynthesis was established by the genetic observations that mutations at this locus eliminate the deposition of these constituents on the surfaces of seedling leaves (Hayes and Brewbaker, 1928; Bianchi, 1975). The identification of *Gl2-like* is solely based on sequence homology with the GL2 and CER2 proteins, without any functional data to support its role in epicuticular lipid biosynthesis. Mutations at the *cer2* locus

appear to block the normal accumulation of epicuticular lipids, by apparently blocking the conversion of C26 and/or C28 fatty acid to C30 fatty acid (Mcnevin et al., 1991; Jenks et al., 1995).

Our transgenic studies established the functionality of *Gl2-like* in epicuticular lipid biosynthesis. Specifically, as with the transgenic expression of *CER2*, the transgenic expression of either *Gl2* or *Gl2-like* restores the normal levels of epicuticular lipids, and restores ability to convert the C26/C28 fatty acids to C30 fatty acid. Therefore, both *Gl2* and *Gl2-like* are functional homologs of *CER2*. However, there are distinct differences induced by the transgenic expression of *Gl2* as compared to *Gl2-like*, which suggest that the two gene products have acquired distinct functionalities since the gene duplication event that gave rise to the two homologs. These differences between *Gl2* and *Gl2-like* are revealed as differences in the ability to complement the *eceriferum* phenotype on Arabidopsis stems, differences in the epicuticular crystalloid morphologies, differences in total epicuticular lipid loads, differences in the epicuticular compositions and intracellular lipidomes in the *cer2-5* lines complemented by each maize gene.

In all these comparisons, the *cer2-5* plants that were genetically complemented with the *Gl2-like* transgene presented traits that are more like the wild-type plants than the *Gl2*-complement plants. Moreover, as with the transgenic expression with the *CER2* gene, the transgenic expression of *Gl2-like* in the wild-type background did not induce any new epicuticular lipid associated traits. In contrast, the parallel experiments conducted with the transgenic expression of *Gl2*, resulted in plants that express new epicuticular lipid associated traits that are not associated with the wild-type host. Collectively, these observations are interpreted to indicate that although both maize *Gl2* and *Gl2-like* genes express a functionality that can replace the *CER2* function of Arabidopsis, the *Gl2-like* gene more completely and accurately recapitulates the *CER2* functionality, whereas *Gl2* has functionalities that are beyond those of *Gl2-like* and *CER2*. These differential

neofunctionalities (i.e., the ability to support the generation of C32 fatty acids and alkyl derivatives) probably arose following the gene duplication event that generated the two maize paralogs. Moreover, this neofunctionalization that enables *Gl*2 to induce the production of C32 alkyl-chains in Arabidopsis, is consistent with the fact that the maize epicuticular lipids are predominantly alkyl chains derived from C32 fatty acids.

Relationship between BAHD catalytic activity and the ability of *Glossy2* and *Glossy2-like* to support *in planta* epicuticular lipid deposition

The BAHD-family of acytransferases are recognizable by the conservation of two primary sequence motifs: a) the HXXXDX catalytic domain; and b) the DFGWG domain that appears to stabilize the substrate-enzyme complex (St-Pierre and Luca, 2000; D'Auria, 2006). Although both these domains have been reported to be essential for enzyme functionality, the latter domain is not fully conserved in all characterized BAHD enzymes (Suzuki et al., 2003; Bayer et al., 2004; Ma et al., 2005; Unno et al., 2007). CER2 and GL2 are archetypal of this class of enzymes, although this was not recognized when they were initially molecularly identified because at that time they represented novel sequences (Tacke et al., 1995; Xia et al., 1996). The identification of the BAHD enzyme family was enabled by the biochemical-based characterization of a number different enzymes that catalyze acyltransferase reactions in the biosynthesis of diverse specialized metabolites (St-Pierre and Luca, 2000; D'Auria, 2006).

We tested whether GL2 and GL2-LIKE require a functional HXXXDX catalytic domain in order to support epicuticular lipid deposition *in planta*. This was done by individually replacing the conserved His and Asp residues, and the terminal X-residue of the domain (Ile) to Ala. The resulting variant enzymes were transgenically expressed in *cer2-5* mutant plants and the stems of the resulting plants were evaluated to assess the accumulation and composition of epicuticular

lipids. These point mutations that are predicted to disrupt BAHD catalytic activity did not affect the ability of GL2 to support epicuticular lipid deposition, whereas these mutations partially disrupted the ability of GL2-LIKE to support this metabolic outcome. These findings establish that GL2 does not require BAHD catalytic activity to support epicuticular lipid deposition, whereas this domain is required in order for GL2-LIKE to achieve the same function.

The GL2 results agree with studies previously conducted with CER2 (Haslam et al., 2012). Specifically, Haslam et al (2012) showed that the transgenic expression of the His-Ala variant in the BAHD-defining HXXXDX domain of CER2 could still complement the *cer2* associated epicuticular lipid deficiency of Arabidopsis stems (i.e., a functional HXXXDX catalytic domain is not needed for *CER2*-function). Moreover, they also showed that the yeast co-expression of CER2 with a 3-ketoacyl-CoA synthase (KCS) component of the plant FAE system, resulted in the ability of the yeast strain to produce longer-chain fatty acids, of 28 and 30 carbon chain length; yeast strains normally only produce C26 VLCFA (Haslam et al., 2012). These results have been interpreted to indicate that CER2 interacts with the FAE system for VLCFA biosynthesis, and alters the product profile of the system. Furthermore, this capability does not require the BAHD acyltransferase catalytic activity of CER2, as mutations in the HXXXDX catalytic does not affect the plant epicuticular lipid profile. Similar conclusions have been reached with the characterization of Arabidopsis *CER2* homologs, such as *CER2-LIKE1* and *CER2-LIKE2*, where the former lacks the His residue in the predicted HXXXDX motif (Haslam et al., 2012, 2015).

In contrast to *Gl2, Gl2-like* needs the intact, and presumably functional BAHD-defining HXXXDX catalytic domain in order to fully complement the *in planta cer2* function in the *cer2-5* mutant. Thus, this is another distinguishing feature between *Gl2* and *Gl2-like* that represents a neofunctionalization that has followed the gene duplication event that generated the two maize

homologs. Moreover, a functional HXXXDX BAHD-catalytic domain appears to also be required by the rice *CER2* homolog, *WSL4* (Wang et al., 2017), which facilitates the interaction with plant KCS to affect fatty acid elongation beyond C24 chain-length.

These characterizations raise a number of questions concerning the ability of these proteins to support epicuticular lipid deposition and the function of other Clade II BAHD proteins: 1) is GL2-LIKE (and possibly the rice homolog, *WSL4*) a bifunctional protein, one mediated by the HXXXDX catalytic function, and the second that is independent of this catalytic function; 2) do other members of the Clade-II BAHD family possess the acyltransferase catalytic activity; and 3) what are the two substrates (the acyl-donor and acyl-acceptor) of the GL2-LIKE acyltransferase enzyme. Answering these questions will probably require the development of a direct biochemical assay that go beyond the studies to date, which primarily rely on correlations between genetic modifications and metabolic outcomes.

The role of *Glossy2* and *Glossy2-like* in supporting the accumulation of VLCFA components of intracellular lipid pools

Mutations in the *cer2* locus appear to block the fatty acid elongation process from C26 or C28 to C30 fatty acids (McNevin et al., 1993; Jenks et al., 1995). Based on our current understanding of the mechanism of fatty acid elongation, it is unclear how specific iterations of the elongation cycle (i.e., from C26 or C28 to C30) can be distinguished from any of the other 6 iterations that elongate a C18 fatty acid to a C30 fatty acid. Recent studies with the Arabidopsis and rice CER2 homologs, have proposed the possibility of a physical interaction between the condensing enzyme of the FAE complex (specifically KCS6) and CER2 (Haslam et al., 2015; Wang et al., 2017) and this interaction may affect a single iteration of the elongation cycles. Based on the epicuticular lipid profiles in transgenic Arabidopsis lines we surmised that as with *CER2,*

Glossy2-like can support the ability of overcoming the elongation of fatty acids from C28 to C30. Although *Glossy2* also shares this capability, it can additionally contribute to the elongation of C30 to C32 fatty acids. Thus, we hypothesize that *Gl2-like* and *Gl2* probably have overlapping functions in maize, where the former affects fatty acid elongation from C26/C28 to C30, and the latter from C26/C28 to C32.

The primary genetic lesion that determines these extracellular epicuticular lipid traits occur in the intracellular lipid metabolic processes that underlie the origins of the epicuticular lipids. These processes are primarily associated with the intracellular membranes (ER and possibly plasma membrane) that house the biochemical reactions associated with VLCFA biosynthesis and down-stream reactions (Bernard and Joubès, 2013). Furthermore, VLCFA metabolism is also juxtaposed with other lipid metabolic pathways, including phospholipids, sphingolipids, storage lipids, and suberin (Li-Beisson et al., 2010). Additionally, the epicuticular lipids are closely associated and integrated with the polymeric cutin matrix, which is assembled by covalent ester bonds (Kolattukudy et al., 1976), the types of bonds that BAHD enzymes generate. Therefore, we profiled and compared the intracellular lipidomic pool and of the cutin matrix of the stems of the different genotypes generated in this study. These comparisons indicate that genetic manipulations associated with mutating the *cer2* function and replacing it with either *Gl2* or *Gl2-like*, not only affect epicuticular lipid profiles, but they also have pronounced effects on the intracellular lipids and cutin monomers, particularly those lipids that utilize VLCFAs, and thus have undergone fatty acid elongation via the ER localized FAE system (Li-Beisson et al., 2010; Bernard and Joubès, 2013). In addition, *CER2, Gl2,* and *Gl2-like* altered the accumulation of several unidentified intracellular lipid molecules. In combination therefore, these data may imply that these genes have regulatory roles in controlling the FAE complex.

Based on the lipidomics analyses and the HXXXDX-mutagenesis studies, it's possible to hypothesize that the acyltransferase reaction of GL2-LIKE may be part of the regulatory mechanism that modulates the FAE complex. However, the mechanism by which such an acyltransferase reaction occurs remains to be determined. It's possible for example to speculate that this BAHD activity acylates the side chain of an amino acid, which can act as a nucleophile (e.g., Cys, Lys, Arg, Asp, Glu, His, Tyr, and Ser) (Hermanson 2008; Bischoff and Schlüter, 2012), and this modification, possibly of a KCS-component, regulates specific iterations of the FAE elongation cycles.

Material and Methods

Plant Material and Growth Conditions

T-DNA mutant line SALK_084443C (*cer2-5*; At4g24510) in the Col-0 background was obtained from the Arabidopsis Biological Resource Center (www.arabidopsis.org). This T-DNA insertion disrupts the second exon of *CER2*. This *cer2-5* mutant and the wild-type Col-0 were used in all the experiments described herein. Seeds were sown in LC1 Sunshine Mix (Sun Gro Horticulture, Bellevue, WA) and treated at 4°C for 5d to break seed dormancy. Seedlings were transferred to individual pots and grown to maturity under constant growth conditions in a regulated growth room at 22°C under continuous illumination (2568 Lux or photosynthetic photon flux density of 100 μmol of photons m^{-2} sec^{-1}). Plants were watered once a week and the irradiance, temperature, and relative humidity were monitored using an Onset Computer Corporation (Bourne, MA) Hobo monitor U12-012 (www.onsetcomp.com). Biochemical and microscopic analyses were conducted on stem tissues of Arabidopsis plants of different genotypes, harvested when the primary flower bolt was 35 to 40-cm in height. Flowers, cauline leaves and siliques were removed and stem tissue was used for analysis.

Maize *glossy2* mutant seeds (glossy2-Salamini; Maize Genetics COOP Stock Center catalog #208H; maizecoop.cropsci.uiuc.edu) were out-crossed to inbred B73 and the resulting heterozygous F1 seeds were selfed and back-crossed to the inbred B73. The F2 seeds were planted in soil, and grown in a climate-controlled greenhouse under a diurnal cycle of 16-h light and 8-h dark at 27 ºC and 24 ºC respectively, maintaining a 30% humidity. The *glossy2* mutant seedlings were identified from the segregating progeny by their water beading phenotype, and non-beading sibling seedlings were used as wild-type controls. Five individual maize seedling plants, at the 3-5 leaf stage were pooled to generate a single replicate, and a total of 4 to 5 replicates where used for epicuticular lipid and lipidomic analyses.

Molecular Cloning

The full-length *Glossy2* (GRMZM2G098239) and *Glossy2-like* (GRMZM2G315767) ORFs were codon optimized for expression in Arabidopsis with GeneOptimizer (GeneArt, LifeTechnologies) and OptimumGene™ (GenScript, Piscataway, NJ; www.genscprit.com), respectively. The *Glossy2* sequence was cloned into pMA-RQ and pDONR221 entry vectors by Life Technologies Corporation (Carlsbad, CA, USA). *Glossy2-like* sequence was initially obtained from GenScript as a pUC57-clone and was sub-cloned into pENTR™/D-TOPO® entry vector (Invitrogen). Further sub-cloning of *Glossy2* and *Glossy2-like* sequences for plant expression experiments were performed using LR Clonase II Enzyme Mix (Invitrogen), using pEarleyGate100 vector (Earley et al., 2006), which controls expression of the transgene with the CaMV 35S promoter. The resultant recombinant vectors (p35S::*ZmGlossy2* and p35S::*ZmGlossy2-like*) were introduced into *Agrobacterium tumefaciens* (strain C58C1) by electroporation (Sambrook et al., 1989). For site-directed mutagenesis experiments, mutation of the His-, Asp-, and Ile- residues of the HXXXDX motif of GL2 and GL2-LIKE were generated

using QuikChange XL Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA). The GLOSSY2 protein was also expressed in *E. coli* BL21AI strain (Invitrogen, Carlsbad, CA, USA, www.thermofisher.com) using the pDEST17 expression vector (Invitrogen). The authenticity of all recombinant plasmids was confirmed by DNA sequencing.

Plant Transformation and Selection

Glossy2 and *Glossy2-like* transgenes were transformed into wild-type Arabidopsis, Col-0 ecotype, and the homozygous *cer2-5* mutant line, which is also in the Col-0 ecotype background, via an Agrobacterium-mediated floral-dip protocol, adapted from Clough and Bent, 1998. Briefly, inflorescence bolts were submerged for 20s in an infiltration medium containing *Agrobacterium tumefaciens* (strain C58C1) carrying either p35S:*ZmGlossy2* or p35S::*ZmGlossy2-like.* The infiltration medium consisted of 2% sucrose and 0.02% Vac-in-stuff Silwet L-77 (Lehle Seeds, Round Rock, TX). Plants were then returned to a 22 ^{o}C growth chamber under continuous illumination until seed-set. Seeds from transformed plants were collected, and germinated in soil. Transformant seedlings were initially identified as being resistant to BASTA herbicide (i.e., glufosinate), applied at a dilution of 1:1000. The herbicide resistant plants were propagated and selfed to the T2 generation. Plants carrying the *Glossy2* or *Glossy2-like* transgene were molecularly confirmed with gene-specific by PCR analysis using DNA-templates isolated from individual herbicide resistant plants. RNA expression of *Glossy2-like* was confirmed by RT-PCR and ubiquitin mRNA (At4g05320) was used as the internal control, GLOSSY2 protein expression was evaluated by Western blot assays. Three replicate lines from each of three independent transformation events were maintained and used for all further experiments.

Western blot analysis

Protein-extracts were prepared by homogenizing plant leaf tissue with a buffer consisting of 62.5 mM Tris-HCl, pH 6.8, 30% glycerol, 10% SDS and 10% 2-mercaptoethanol. Samples were vortexed for 5-min, boiled for 10-min and centrifuged at 13000g for 2-min. The clear supernatant protein extracts were subjected to SDS-PAGE and proteins were transferred to a nitrocellulose membrane according to manufacturer's instructions (Bio-RAD, Hercules, CA). Protein blots were first probed with the GLOSSY2-specific antiserum (1:1000 dilution), and then with horseradish peroxidase-linked anti-mouse IgG antibody (Bio-Rad) (1:1000 dilution). The antigen-antibody complexes were detected using the Pierce ECL chemi-luminescent detection system (Thermo Scientific, Rockford, IL) and visualized on the ChemiDoc XRS+ gel documentation system (Bio-Rad).

Stereomicroscopy

The *eceriferum* phenotype on stems of plants was visualized in the central 1-cm segment of the stems using a Zeiss macro-Zoom microscope (Zeiss Axio Zoom V16) with ZEN2 software (Carl Zeiss Inc., Thornwood, NY).

Scanning Electron Microscopy

The 1-cm-long stem segments were mounted on aluminum stubs with double-sided carbon tape, dried in a desiccator and sputter-coated (www.tedpella.com) with a Cressington HR208 sputter coater with platinum for 90s at 40 milliamps, depositing a 10 nm-thick coating. The segments were mounted on aluminum stubs, examined at 10 kV with a Hitachi SU4800 field emission SEM (www.hitachi-hightech.com), and images were digitally captured in TIFF format.

Extraction and analysis of stem extracellular epicuticular lipids

Extracellular epicuticular lipids were extracted from isolated stem tissue by immersing the stems for 10-seconds in to 10-mL chloroform (HPLC-grade Fisher Chemicals, Pittsburgh, PA), containing 5-µg hexacosane (Sigma-Aldrich, Milwaukee, WI) as an internal standard. The stems were then flash frozen using liquid nitrogen, lyophilized (Labconco FreeZone Benchtop Freeze Dry System), and the weight of the dry biomass was determined. A similar extraction procedure was used to extract epicuticular lipids from maize seedling, collected at 3-5 leaf stage, using 60-µg hexacosane as the internal standard.

The chloroform was removed from the extracellular epicuticular lipid extracts by evaporation under a stream of N_2 gas, and the dried lipids were silylated using a protocol based on that of Wood et al., (2001) and Hannoufa et al., (1993). Specifically the nitrogen-dried extracts were dissolved in 0.2-mL acetonitrile and silylation was performed by the addition of 0.05-mL of *N*,*O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS) (Sigma-Aldrich), and incubated at 65°C for 30 min. The samples were cooled, dried under a stream of N_2 gas and resuspended in chloroform. One-microliter of the derivatized sample was injected into GC-MS or GC-FID.

GC-MS analysis was conducted with an Agilent 7890A GC, equipped with a HP-5ms capillary column (30 m x 0.25 mm x 0.25 µm, Agilent), interfaced to an Agilent 5975C quadrapole MS. Chromatography was conducted with helium gas, at a flow-rate of 1.0 mL/min, and an inlet temperature at 280^{o}C. The column oven temperature was initially held at 120°C, then ramped at 10°C/min to 260°C and held at this temperature for 10 min, and then ramped at 5°C/min to 320°C and held there for 4 min. EI-MS ionization energy was set to 70 eV and the interface temperature was 280°C. Resulting chromatograms and mass-spectra were deconvoluted and queried against an

in-house Mass-Spectral library and the NIST 14 Mass Spectral Library using the NIST AMDIS software (Stein, 1999).

GC-FID analysis was conducted with an Agilent 6890 GC, equipped with a DB-1 MS capillary column (15 m x 0.25 mm x 0.25 μm, Agilent 122-0112). Chromatography was conducted with helium gas, at a flow-rate of 1.2 mL/min, and an inlet temperature at 280^{o}C. The column oven temperature was initially held at 80°C, then ramped at 15°C/min to 220°C, then ramped at 7.5°C/min to 310°C, and finally ramped at 20°C/min to 340°C and held there for 5 min. For peak identification purposes chromatograms were compared with parallel GC-MS analysis.

The relative abundance of extracellular lipids/mg dry weight of plant material was calculated based on the ion-strength of the internal standard. Statistical significance was determined using Student's t-test.

Extraction and analysis of stem cutin monomers

Cutin was extracted from 50-100 mg of fresh harvested stems tissue. Tissue was initially delipidated by immersing in preheated (85^{o}C) isopropanol containing 0.01% BHT for 10 min and vortexed overnight at room temperature. The following day the solvent was discarded and tissue was re-extracted once with isopropanol containing 0.01% BHT, twice with chloroform:methanol (2:1) containing 0.01% BHT, and lastly, once with methanol containing 0.01% BHT, for a period of 2-h each at room temperature. Finally, the methanol was discarded and the residue tissue was dried in a lyophilizer to constant dry weight.

Cutin depolymerization was performed using acid-catalyzed transmethylation on the dried delipidated tissue using 2-mL freshly prepared methanolic sulfuric acid (4%) and 0.2-mL toluene, spiked with 10-μg heptadecanoic acid as an internal standard. The mixture was heated at 80^{o}C for 2-h, cooled, and extracted with 4-mL dichloromethane and 1-mL 0.9% NaCl (w/v) in 100 mM

Tris-HCl buffer, pH 8.0. The mixture was vortexed for 10-min and centrifuged at 1500 rpm for 2-min to separate the polar from the non-polar phases. The non-polar phase was collected, dried, and silylated. Silylation was performed at 110^{o}C for 10-min with 100 μL anhydrous pyridine and 100 μL BSTFA [*N,O-bis*(trimethylsilyl)-trifluoroacetamide]. Upon drying under a gentle stream of nitrogen gas, the residue was dissolved in 500 μL of heptane:toluene (1:1, v/v) and subjected to GC-MS analysis.

Samples were analyzed using an Agilent 7890A GC, equipped with a HP-5ms capillary column (30 m x 0.25 mm x 0.25 μm, Agilent) and interfaced to an Agilent 5975C quadrapole MS. Chromatography was conducted with helium gas, at a flow-rate of 2.0 mL/min, and an inlet temperature at 280^{o}C. The column oven temperature was initially held at 110°C, then ramped at 10°C/min to 300°C and held at this temperature for 4 min, and then ramped at 5°C/min to 320°C and held there for 4 min . Resulting chromatograms were analyzed and previously mentioned.

Extraction and analysis of stem lipidome

Intracellular lipids were extracted from 5-8 mg dry Arabidopsis stem tissues or dried maize seedling leaf tissues that had been previously been extracted for epicuticular lipids. The samples were analyzed using a Waters Xevo G2 Q-TOF MS equipped with a Waters ACQUITY UPLC system as previously reported by Okazaki et al., 2015. Briefly, flash frozen plant tissue was lyophilized and pulverized to fine powder using a Mixer Mill 301 (Retsch GmbH, Germany). An extraction solvent containing a mixture of methyl tert-butyl ether and methanol (3:1, v/v), spiked with 1 μM 1,2-didecanoyl-sn-glycero-3-phosphocholine (Sigma-Aldrich) as internal standard, was added to the tissue, at a rate of 160 μl per mg dry tissue and the mixture was thoroughly vortexed. Water was added (50-μl per mg tissue) to the extract and the mixture was thoroughly mixed for 10-min at room temperature using a sample tube mixer. The mixture was incubated on ice for 10-

min and centrifuged at 3000 g at 4^{o}C for 10 min. 160 µl of the supernatant was collected and evaporated to dryness in a SpeedVac. The residue was dissolved in 200 µl of ethanol, vortexed for 10-min at room temperature, and centrifuged at 10,000 g at 4^{o}C for 10-min. The supernatant (180-µl) was used immediately for lipid analysis.

Lipid analysis was conducted using liquid chromatography quadrupole time-of-flight mass spectrometer (HPLC, Waters Acquity UPLC system; MS, Waters Xevo G2 Qtof), as reported by Kimbara et al., 2013. Briefly, a two-solvent (Solvents A and B) system was used to separate metabolites. Solvent A comprised of acetonitrile:water:1 M ammonium acetate:formic acid (158 g: 800 g: 10 mL: 1 mL); and solvent B comprised of acetonitrile:2-propanol:1 M ammonium acetate:formic acid (79 g: 711 g: 10 mL:1 mL). Chromatography was with an HPLC column, Acquity UPLC HSS T3 (pore size, 1.8 µm; 1.0 i.d × 50 mm long; Waters), and gradient program was 35% B at 0 min, 70% B at 3 min, 85% B at 7 min, 90% B at 10 min, 90% B at 12 min and 35% B at 12.5 min, at a flow rate of 0.15 ml min^{-1}; temperature, 55°C; MS detection. The conditions for recording MS/MS spectra in positive-ion mode were as follows: scan range, m/z 100–1200; capillary, 3.0 kV; sampling cone, 20V; extraction cone, 4.0V; source temperature, 120 °C; desolvation temperature, 450 °C; cone gas flow, 50 l h^{-1}; desolvation gas flow, 600 l h^{-1}; collision energy started at 20 eVand was ramped to 50 eV. The conditions for recording MS/MS spectra in negative-ion mode were as follows: scan range, m/z 100–1200; capillary, 2.5 kV; sampling scone, 40V; extraction cone, 4.0V; source temperature, 120 °C; desolvation temperature, 450 °C; cone gas flow, 50 l h^{-1}; desolvation gas flow, 600 l h^{-1}; collision energy started at 20 eV and was ramped to 50 eV.

LCMS data matrices recorded in the MassLynx4.1 format (raw) were processed using Markerlynx XS (Waters). The data matrices were searched against an in-house lipid MS library

(RIKEN, Japan). To normalize the data, the peak intensity values of each ion were divided by the internal standard ion (*m/z* 566.382 $[M + H]^+$) for positive ion mode, and the ion (*m/z* 610.372 $[M + HCOO]^-$) for negative ion mode.

Acknowledgements

The authors acknowledge Dr. Harry T. Horner, Tracey Stewart and Randell Den Adel for helpful discussion and for use of equipment at Roy J. Carver High Resolution Microscopy Facility (Iowa State University, Ames, IA); Drs. Zhihong Song, Ann Perera and Lucas Showman of the WM Keck Metabolomics Research Laboratory (Iowa State University, Ames, IA) for assistance in metabolic analyses; Dr. Marna-Yandeau Nelson and Bri Vidrine (Iowa State University) for providing maize *glossy2* mutant seed, and propagating this genetic stock. Hybridoma Facility (Iowa State University, Ames, IA) for characterization of GLOSSY2 antibody. This work was partly funded by grants from the National Science Foundation through awards IOS-1139489 and EEC-0813570 to BJN, and an EAPSI award OISE-1614020 to LEA, and by the State of Iowa through the Center for Metabolic Biology.

References

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C., Ecker, J.R., 2003. Genome-Wide Insertional Mutagenesis of Arabidopsis thaliana. Science. 301, 653 LP-657. https://doi.org/10.1126/science.1086391

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25. https://doi.org/10.1093/nar/25.17.3389

Andorf, C.M., Cannon, E.K., Portwood 2nd, J.L., Gardiner, J.M., Harper, L.C., Schaeffer, M.L., Braun, B.L., Campbell, D.A., Vinnakota, A.G., Sribalusu, V. V, Huerta, M., Cho, K.T., Wimalanathan, K., Richter, J.D., Mauch, E.D., Rao, B.S., Birkett, S.M., Sen, T.Z., Lawrence-

Dill, C.J., 2016. MaizeGDB update: new tools, data and interface for the maize model organism database. Nucleic Acids Res. 44, D1195–D1201. https://doi.org/10.1093/nar/gkv1007

Bayer, A., Ma, X., Stöckigt, J., 2004. Acetyltransfer in natural product biosynthesis - Functional cloning and molecular analysis of vinorine synthase. Bioorganic Med. Chem. 12, 2787–2795. https://doi.org/10.1016/j.bmc.2004.02.029

Bernard, A., Joubès, J., 2013. Arabidopsis cuticular waxes: Advances in synthesis, export and regulation. Prog. Lipid Res. 52, 110–129. https://doi.org/10.1016/j.plipres.2012.10.002

Bianchi, 1975. Glossy mutants of maize. VI. Chemical constituents of glossy-2 epicuticular waxes. Maydica 20, 165–173.

Bianchi Rome (Italy)), A. (Istituto S. per la C., Bianchi, G., Avato, P. (Pavia U. (Italy). D. di C.O., Salamini Koeln (Germany, F.R.). Erwin-Baur-Inst.), F. (Max-P.I. fuer Z., 1985. Biosynthetic pathways of epicuticular wax of maize as assessed by mutation, light, plant age and inhibitor studies. Maydica (Italy).

Bischoff, R., Schlüter, H., 2012. Amino acids : Chemistry , functionality and selected non-enzymatic post-translational modifications ☆ Amino group. J. Proteomics 75, 2275–2296. https://doi.org/10.1016/j.jprot.2012.01.041

Bonaventure, G., Beisson, F., Ohlrogge, J., Pollard, M., 2004. Analysis of the aliphatic monomer composition of polyesters associated with Arabidopsis epidermis: Occurrence of octadeca-cis-6, cis-9-diene-1,18-dioate as the major component. Plant J. 40, 920–930. https://doi.org/10.1111/j.1365-313X.2004.02258.x

Clough, S.J., Bent, A.F., 1999. Floral dip : a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana 16, 735–743.

D'Auria, J.C., 2006. Acyltransferases in plants: a good time to be BAHD. Curr. Opin. Plant Biol. 9, 331–340. https://doi.org/10.1016/j.pbi.2006.03.016

Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., Pikaard, C.S., 2006. Gateway-compatible vectors for plant functional genomics and proteomics. Plant J. 45, 616–629. https://doi.org/10.1111/j.1365-313X.2005.02617.x

Fernandes, A.M.S., Baker, E.A., Martin, J.T., 1964. Studies on plant cuticle. Ann. Appl. Biol. 53, 43–58. https://doi.org/10.1111/j.1744-7348.1964.tb03779.x

Franke, R., Briesen, I., Wojciechowski, T., Faust, A., Yephremov, A., Nawrath, C., Schreiber, L., 2005. Apoplastic polyesters in Arabidopsis surface tissues - A typical suberin and a particular cutin. Phytochemistry 66, 2643–2658. https://doi.org/10.1016/j.phytochem.2005.09.027

Garvey, G.S., Mccormick, S.P., Rayment, I., 2008. Structural and Functional Characterization of the TRI101 Trichothecene 3- O -Acetyltransferase from Fusarium sporotrichioides and Fusarium graminearum 283, 1660–1669. https://doi.org/10.1074/jbc.M705752200

Hannoufa, A., McNevin, J., Lemieux, B., 1993. Epicuticular waxes of eceriferum mutants of Arabidopsis thaliana. Phytochemistry 33, 851–855. https://doi.org/https://doi.org/10.1016/0031-9422(93)85289-4

Haslam, T.M., Haslam, R., Thoraval, D., Pascal, S., Delude, C., Domergue, F., Fernández, A.M., Beaudoin, F., Napier, J.A., Kunst, L., Joubès, J., 2015. ECERIFERUM2-LIKE Proteins Have Unique Biochemical and Physiological Functions in Very-Long-Chain Fatty Acid Elongation. Plant Physiol. 167, 682–92. https://doi.org/10.1104/pp.114.253195

Haslam, T.M., Mañas-Fernández, A., Zhao, L., Kunst, L., 2012. Arabidopsis ECERIFERUM2 is a component of the fatty acid elongation machinery required for fatty acid extension to exceptional lengths. Plant Physiol. 160, 1164–74. https://doi.org/10.1104/pp.112.201640

Hayes, H.K., Brewbaker, H.E., 1928. Glossy Seedlings in Maize. Am. Nat. 62, 228–235.

Heredia, A., 2003. Biophysical and biochemical characteristics of cutin, a plant barrier biopolymer. Biochim. Biophys. Acta - Gen. Subj. 1620, 1–7. https://doi.org/https://doi.org/10.1016/S0304-4165(02)00510-X

Hermanson, G.T., 2008. Bioconjugate techniques (2nd ed.), Academic Press, London

Jellings, A.J., Leech, R.M., 1982. The Importance of Quantitative Anatomy in the Interpretation of Whole Leaf Biochemistry in Species of Triticum, hordeum and Avena. New Phytol. 92, 39–48.

Jenks, M. A., Tuttle, H. A., Eigenbrode, S.D., Feldmann, K. A., 1995. Leaf Epicuticular Waxes of the Eceriferum Mutants in Arabidopsis. Plant Physiol. 108, 369–377. https://doi.org/10.1104/pp.108.1.369

Jenks, M.A., Joly, R.J., Peters, P.J., Rich, P.J., Axtell, J.D., Ashworth, E.N., 1994. Chemically Induced Cuticle Mutation Affecting Epidermal Conductance to Water Vapor and Disease Susceptibility in Sorghum bicolor (L.) Moench. Plant Physiol. 105, 1239–1245. https://doi.org/10.1104/pp.105.4.1239

Jetter, R., Schäffer, S., Riederer, M., 2000. Leaf cuticular waxes are arranged in chemically and mechanically distinct layers: evidence from Prunus laurocerasus L. Plant. Cell Environ. 23, 619–628. https://doi.org/10.1046/j.1365-3040.2000.00581.x

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J.E., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10, 845–858. https://doi.org/10.1038/nprot.2015.053

Kelley, L.A., Sternberg, M.J.E., 2009. Protein structure prediction on the Web : a case study using the Phyre server 4. https://doi.org/10.1038/nprot.2009.2

Kerstiens, G., 1996. Cuticular water permeability and its physiological significance. J. Exp. Bot. 47, 1813–1832.

Kimbara, J., Yoshida, M., Ito, H., Kitagawa, M., Takada, W., Hayashi, K., Shibutani, Y., Kusano, M., Okazaki, Y., Nakabayashi, R., Mori, T., Saito, K., Ariizumi, T., Ezura, H., 2013. Inhibition of cutin deficient 2 causes defects in cuticle function and structure and metabolite changes in tomato fruit. Plant Cell Physiol. 54, 1535–1548. https://doi.org/10.1093/pcp/pct100

Kolattukudy, P.E., 1980. Cutin, suberin and waxes. In PK Stumpf, ed, The Biochemistry of Plants—A Comprehensive Treaty, Vol 4. Academic Press, New York, pp 571–645

Kolattukudy, P.E., 1985. Enzymatic Penetration of the Plant Cuticle by Fungal Pathogens. Annu. Rev. Phytopathol. 23, 223–250. https://doi.org/10.1146/annurev.py.23.090185.001255

Kolattukudy, P.E., 1965. Biosynthesis of Wax in Brassica oleracea *. Biochemistry 4, 1844–1855. https://doi.org/10.1021/bi00885a023

Kolattukudy, P.E., Croteau, R., and Buckner, J.S., 1976. Biochemistry of plant waxes, in P.E. Kolattukudy (ed.), Chemistry and Biochemistry of Natural Waxes, Elsevier Press, New York, pp. 289 - 347.

Koornneef, M., Hanhart, C.J., Thiel, F., 1989. A Genetic and Phenotypic Description of Eceriferum (cer) Mutants in Arabidopsis thaliana. J. Hered. 80, 118–122.

Kurdyukov, S., Faust, A., Nawrath, C., Bär, S., Voisin, D., Efremova, N., Franke, R., Schreiber, L., Saedler, H., Métraux, J.-P., Yephremov, A., 2006. The epidermis-specific extracellular BODYGUARD controls cuticle development and morphogenesis in Arabidopsis. Plant Cell 18, 321–339. https://doi.org/10.1105/tpc.105.036079

Leide, J., Hildebrandt, U., Reussing, K., Riederer, M., Vogg, G., 2007. The Developmental Pattern of Tomato Fruit Wax Accumulation and Its Impact on Cuticular Transpiration Barrier Properties: Effects of a Deficiency in a β-Ketoacyl-Coenzyme A Synthase (LeCER6). Plant Physiol. 144, 1667 LP-1679. https://doi.org/10.1104/pp.107.099481

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M.X., Arondel, V., Bates, P.D., Baud, S., Bird, D., Debono, A., Durrett, T.P., Franke, R.B., Graham, I.A., Katayama, K., Kelly, A.A., Larson, T., Markham, J.E., Miquel, M., Molina, I., Nishida, I., Rowland, O., Samuels, L., Schmid, K.M., Wada, H., Welti, R., Xu, C., Zallot, R., Ohlrogge, J., 2010. Acyl-lipid metabolism. Arab. B. 8, e0133–e0133. https://doi.org/10.1199/tab.0133

Loneman, D.M., Peddicord, L., Al-Rashid, A., Nikolau, B.J., Lauter, N., Yandeau-Nelson, M.D., 2017. A robust and efficient method for the extraction of plant extracellular surface lipids as applied to the analysis of silks and seedling leaves of maize. PLoS One 12, e0180850.

Long, L.M., Patel, H.P., Cory, W.C., Stapleton, A.E., 2003. The maize epicuticular wax layer provides UV protection. Funct. Plant Biol. 30, 75–81.

Ma, X., Koepke, J., Panjikar, S., Fritzsch, G.G., St??ckigt, J., Stöckigt, J., 2005. Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. J. Biol. Chem. 280, 13576–13583. https://doi.org/10.1074/jbc.M414508200

Martin JT, J.B. 1970., 1970. The Cuticles of plants: J.T. Martin and B.E. Juniper. Edward Arnold (Publishers) Ltd., London and New York, pp. x+347, 1970. (Book Review). Micron (1969). https://doi.org/10.1016/0047-7206(70)90020-8

McNevin, J.P., Woodward, W., Hannoufa, A., Feldmann, K.A., and Lemieux, B., 1991. Isolation and characterization of eceriferum (cer) mutants induced by T-DNA insertions in Arabidopsis thaliana.

Molina, I., Bonaventure, G., Ohlrogge, J., Pollard, M., 2006. The lipid polyester composition of Arabidopsis thaliana and Brassica napus seeds. Phytochemistry 67, 2597–2610. https://doi.org/10.1016/j.phytochem.2006.09.011

Moose, S.P., Sisco, P.H., 1996. Glossy15, an APETALA2-like gene from maize that regulates leaf epidermal cell identity. Genes Dev. 10, 3018. https://doi.org/10.1101/gad.10.23.3018

Okazaki, Y., Nishizawa, T., Takano, K., Ohnishi, M., Mimura, T., Saito, K., 2015. Induced accumulation of glucuronosyldiacylglycerol in tomato and soybean under phosphorus deprivation. Physiol. Plant. 155, 33–42. https://doi.org/10.1111/ppl.12334

Pascal, S., Bernard, A., Sorel, M., Pervent, M., Vile, D., Haslam, R.P., Napier, J.A., Lessire, R., Domergue, F., Joubès, J., 2013. The Arabidopsis cer26 mutant, like the cer2 mutant, is specifically affected in the very long chain fatty acid elongation process. Plant J. 73, 733–746. https://doi.org/10.1111/tpj.12060

Perera, M.A.D.N., Qin, W., Yandeau-Nelson, M., Fan, L., Dixon, P., Nikolau, B.J., 2010. Biological origins of normal-chain hydrocarbons: A pathway model based on cuticular wax analyses of maize silks. Plant J. 64, 618–632. https://doi.org/10.1111/j.1365-313X.2010.04355.x

Renny-Byfield, S., Wendel, J.F., 2014. Doubling down on genomes: Polyploidy and crop plants. Am. J. Bot. 101, 1711–1725. https://doi.org/10.3732/ajb.1400119

Riederer, M., 2006. Thermodynamics of the water permeability of plant cuticles : characterization of the polar pathway 1–6. https://doi.org/10.1093/jxb/erl053

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual, 2nd edn. New York: Cold Spring Harbor Laboratory Press

Schnable PS, Stinard PS, Wen TJ, Heinen S, Weber D, Zhang L, Hansen JD, and Nikolau BJ (1994) The genetics of cuticular wax biosynthesis, Maydica 39, 279-287

Sieber, P., Schorderet, M., Ryser, U., Buchala, A., Kolattukudy, P., Métraux, J.-P., Nawrath, C., 2000. Transgenic Arabidopsis Plants Expressing a Fungal Cutinase Show Alterations in the Structure and Properties of the Cuticle and Postgenital Organ Fusions. Plant Cell 12, 721 LP-737. https://doi.org/10.1105/tpc.12.5.721

St-Pierre, B., Luca, V. De, 2000. Chapter Nine Evolution of acyltransferase genes: Origin and diversification fo the BAHD superfamily of acyltransferases involved in secondary metabolism. Recent Adv. Phytochem. 34, 285–315. https://doi.org/10.1016/S0079-9920(00)80010-6

Stein, S.E., 1999. An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. J. Am. Soc. Mass Spectrom. 10, 770–781. https://doi.org/https://doi.org/10.1016/S1044-0305(99)00047-1

Suzuki, H., Nakayama, T., Nishino, T., 2003. Proposed Mechanism and Functional Amino Acid Residues of Malonyl-CoA:Anthocyanin 5-O-Glucoside-6‘ ‘-O-Malonyltransferase from Flowers of Salvia splendens, a Member of the Versatile Plant Acyltransferase Family. Biochemistry 42, 1764–1771. https://doi.org/10.1021/bi020618g

Tacke, E., Korfhage, C., Michel, D., Maddaloni, M., Motto, M., Lanzini, S., Salamini, F., Döring, H.-P., 1995. Transposon tagging of the maize Glossy2 locus with the transposable element En/Spm. Plant J. 8, 907–917. https://doi.org/10.1046/j.1365-313X.1995.8060907.x

Tulloch, A.P. (1976) Biochemistry of plant waxes, in P.E. Kolattukudy (ed.), Chemistry and Biochemistry of Natural Waxes, Elsevier Press, New York, pp. 235 – 287

Unno, H., Ichimaida, F., Suzuki, H., Takahashi, S., Tanaka, Y., Saito, A., Nishino, T., Kusunoki, M., Nakayama, T., 2007. Structural and mutational studies of anthocyanin malonyltransferases establish the features of BAHD enzyme catalysis. J. Biol. Chem. 282, 15812–15822. https://doi.org/10.1074/jbc.M700638200

Vogg, G., Fischer, S., Leide, J., Emmanuel, E., Jetter, R., Levy, A.A., Riederer, M., 2004. Tomato fruit cuticular waxes and their effects on transpiration barrier properties: functional characterization of a mutant deficient in a very-long-chain fatty acid β-ketoacyl-CoA synthase. J. Exp. Bot. 55, 1401–1410.

Wang, X., Guan, Y., Zhang, D., Dong, X., Tian, L., Qu, L.Q., 2017. A β -Ketoacyl-CoA Synthase Is Involved in Rice Leaf Cuticular Wax Synthesis and Requires a CER2-LIKE Protein as a Cofactor. Plant Physiol. 173, 944–955. https://doi.org/10.1104/pp.16.01527

Wood, K. V, Bonham, C.C., Jenks, M.A., 2001. The effect of water on the ion trap analysis of trimethylsilyl derivatives of long-chain fatty acids and alcohols. Rapid Commun. Mass Spectrom. 15, 873–877. https://doi.org/10.1002/rcm.285

Xia, Y., Nikolau, B.J., Schnable, P.S., 1996. Cloning and characterization of CER2, an Arabidopsis gene that affects cuticular wax accumulation. Plant Cell 8, 1291–1304. https://doi.org/10.1105/tpc.8.8.1291

Xu, L., Zeisler, V., Schreiber, L., Gao, J., Hu, K., Wen, J., Yi, B., Shen, J., Ma, C., Tu, J., Fu, T., 2017. Overexpression of the Novel Arabidopsis Gene At5g02890 Alters Inflorescence Stem Wax Composition and Affects Phytohormone Homeostasis. Front. Plant Sci. 8, 68. https://doi.org/10.3389/fpls.2017.00068

Yephremov, A., Wisman, E., Huijser, P., Huijser, C., Wellesen, K., Saedler, H., 1999. Characterization of the FIDDLEHEAD Gene of Arabidopsis Reveals a Link between Adhesion Response and Cell Differentiation in the Epidermis. Plant Cell 11, 2187 LP-2201. https://doi.org/10.1105/tpc.11.11.2187

Yu, D., Ranathunge, K., Huang, H., Pei, Z., Franke, R., Schreiber, L., He, C., 2008. Wax Crystal-Sparse Leaf1 encodes a β–ketoacyl CoA synthase involved in biosynthesis of cuticular waxes on rice leaf. Planta 228, 675–685. https://doi.org/10.1007/s00425-008-0770-9

All supplementary Tables as excel files can be found deposited along with this dissertation at ProQuest under the headings

1) **Chapter 5_Supplemental Table 1_Epicuticular lipid abundance in Arabidopsis and maize leaves**
2) **Chapter 5_Supplemental Table 2_Arabidopsis stems expressing mutated Gl2 and Gl2-like transgenes**
3) **Chapter 5_Supplemental Table 3_Lipidomic and cutin data**

Figures

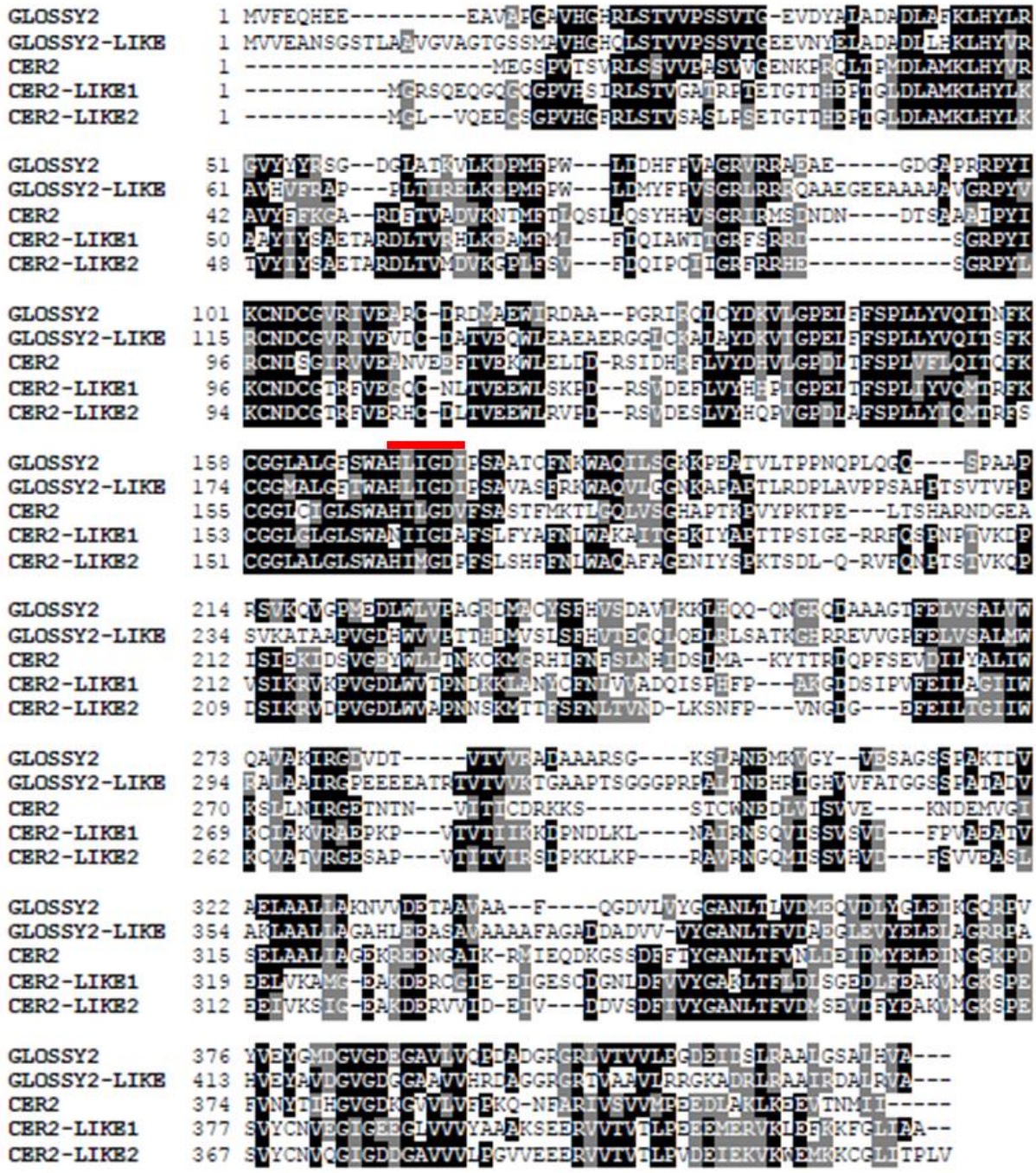


Figure 1. Amino acid sequence comparison of *Zea mays* GLOSSY2 homologs. The maize GLOSSY2 (GRMZM2G098239) and GLOSSY2-LIKE (GRMZM2G315767) sequences are compared to the Arabidopsis CER2 (At4g24510), CER2-LIKE1 (At4g13840) and CER2-LIKE2 (At3g23840) sequences using Clustal O (1.2.4) and BOXSHADE (v3.21). Black shading identifies identical residues, and gray-shading identifies similar residues. The conserved -HXXXDX- acyl-CoA-dependent acyltransferase domain that characterizes BAHD enzymes is identified with a red line above the GL2 sequence.

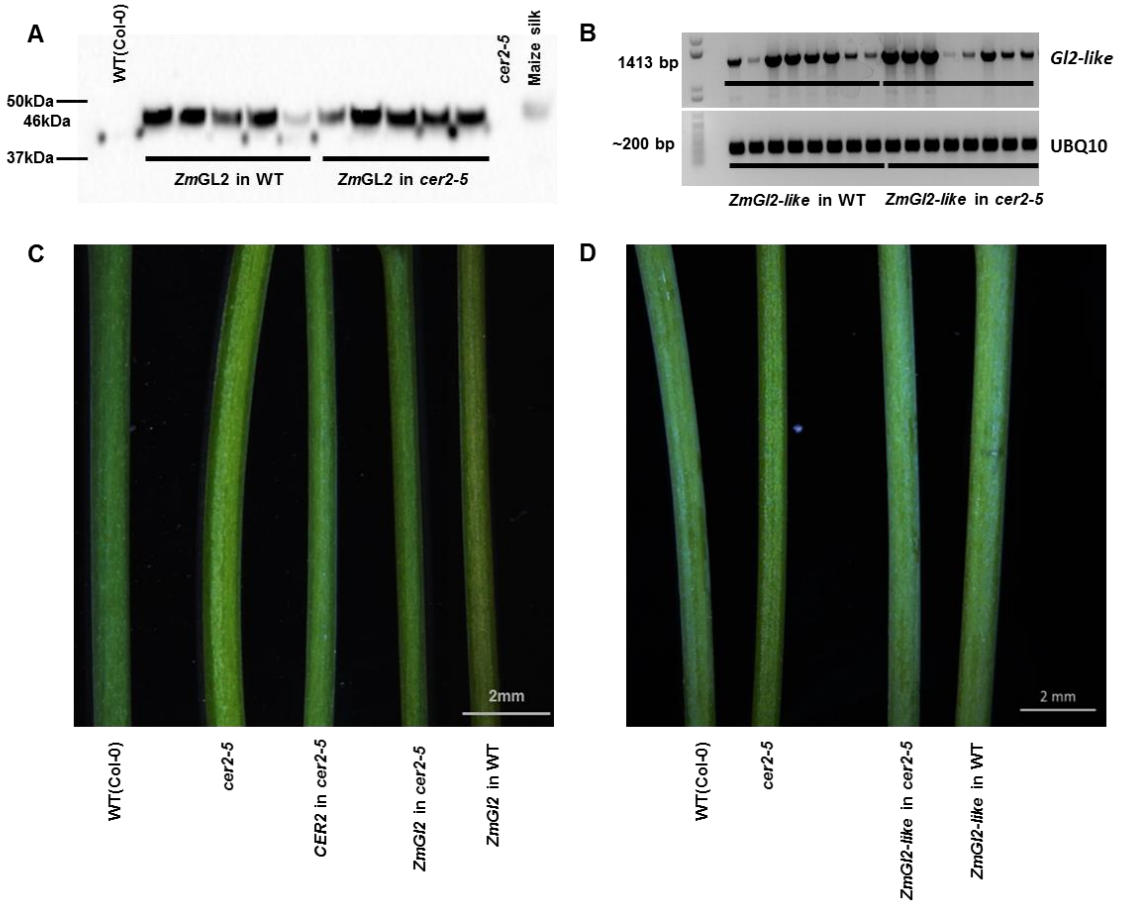


Figure 2. Transgenic expression of *Gl2* and *Gl2-like* in Arabidopsis. A) Western blot analysis of the 46-kDa GLOSSY2 protein transgenically expressed in either wild-type plants or *cer2-5* mutant plants. Positive control is the analysis of protein extracts from maize silks, and negative control is the analysis of protein extracts from stems of wild-type Col-0, and *cer2-5* mutant plants. B) RT-PCR analysis of the *Gl2-like* mRNA transgenically expressed in either wild-type or *cer2-5* mutant plants. Ubiquitin mRNA (At4g05320) is used as the internal control. C) Phenotypes of non-transgenic Arabidopsis wild-type and *cer2-5* mutant stems and transgenic stems expressing either maize *Gl2* or *CER2* in a wild-type and *cer2-5* mutant backgrounds. D) Phenotypes of non-transgenic Arabidopsis wild-type and *cer2-5* mutant stems and transgenic stems expressing the maize *Gl2-like* transgene in either wild-type or *cer2-5* mutant backgrounds.

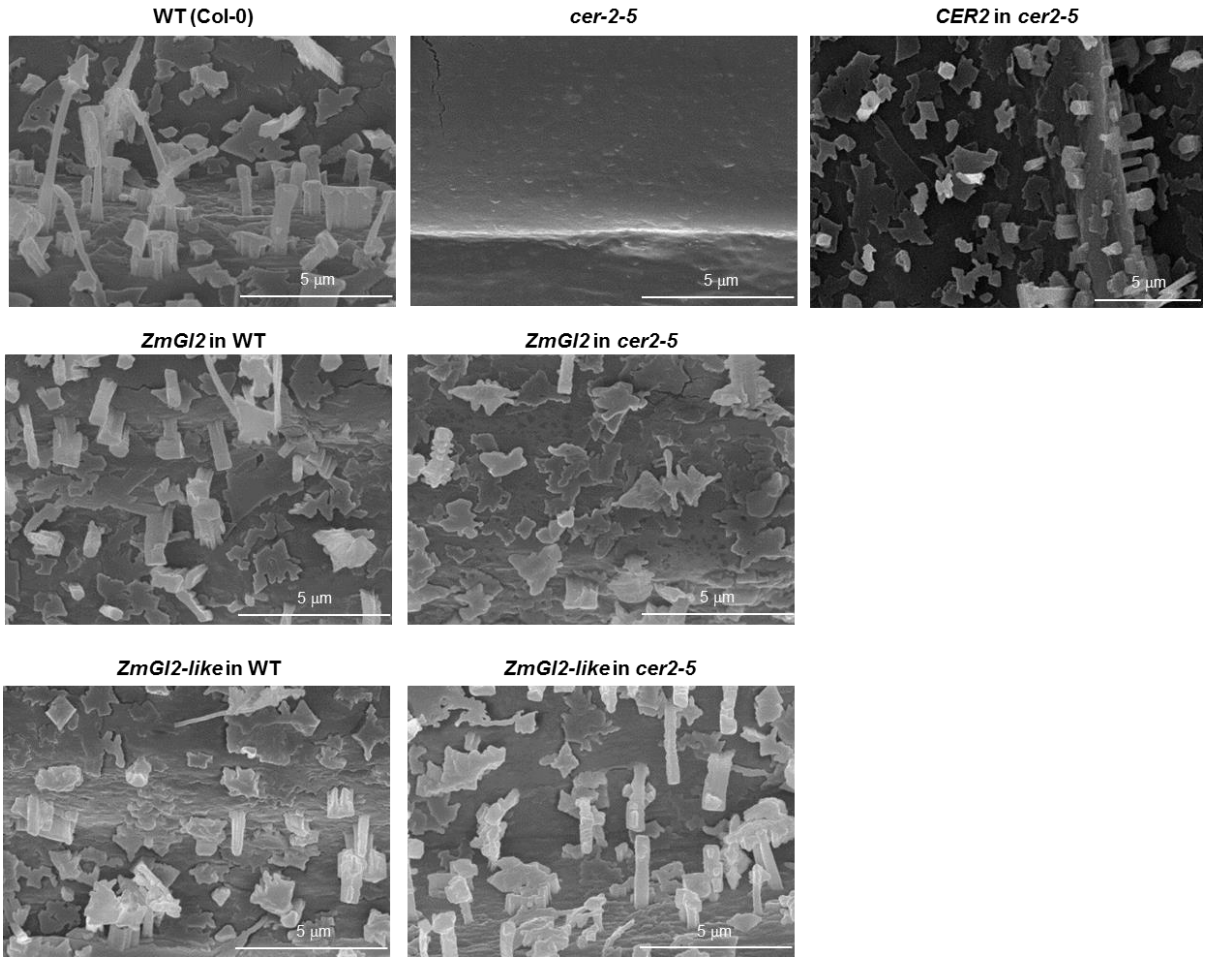


Figure 3. Arabidopsis stem epicuticular crystalloids. Scanning electron micrographs (10000X magnification) of stem surfaces of non-transgenic wild-type or cer2-5 mutant plants, compared to stem surfaces of transgenic plants expressing *Glossy2*, *Glossy2-like* or *CER2* transgenes in either a wild-type or *cer2-5* mutant plants, and *CER2* overexpressed in *cer2-5* mutant background. Scale bars: 5 µm.

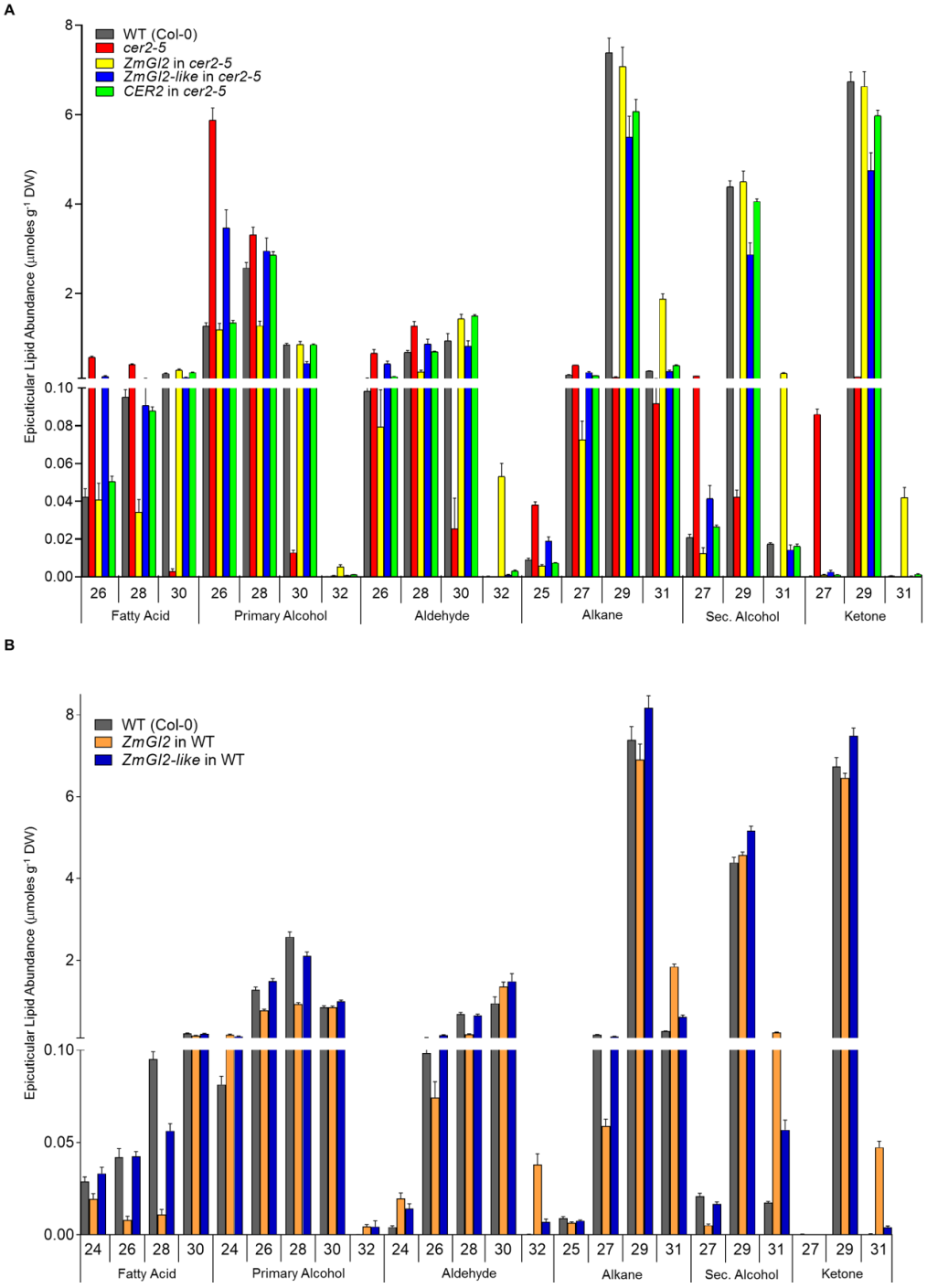


Figure 4. Effect of transgenic expression of *Gl2* and *Gl2-like* on the extracellular epicuticular surface lipid profiles of Arabidopsis stems. A) Transgenic expression of *Gl2* and *Gl2-like* in the *cer2-5* mutant background, as compared to the profiles in the non-transgenic wild-type (Col-0) and *cer2-5* controls. B) Transgenic expression of *Gl2* and *Gl2-like* in the wild-type background. Epicuticular surface lipids were extracted from stems and analyzed as described in the Materials and Methods. Lipids are identified by the carbon chain-length, and by the alkyl chemical class. Data for the minor metabolites of C20-C24 carbon chain-length is in Supplemental Figure 1. All numeric data can be found in Supplemental Table 1. The data represent mean ± standard error of 10-15 replicates

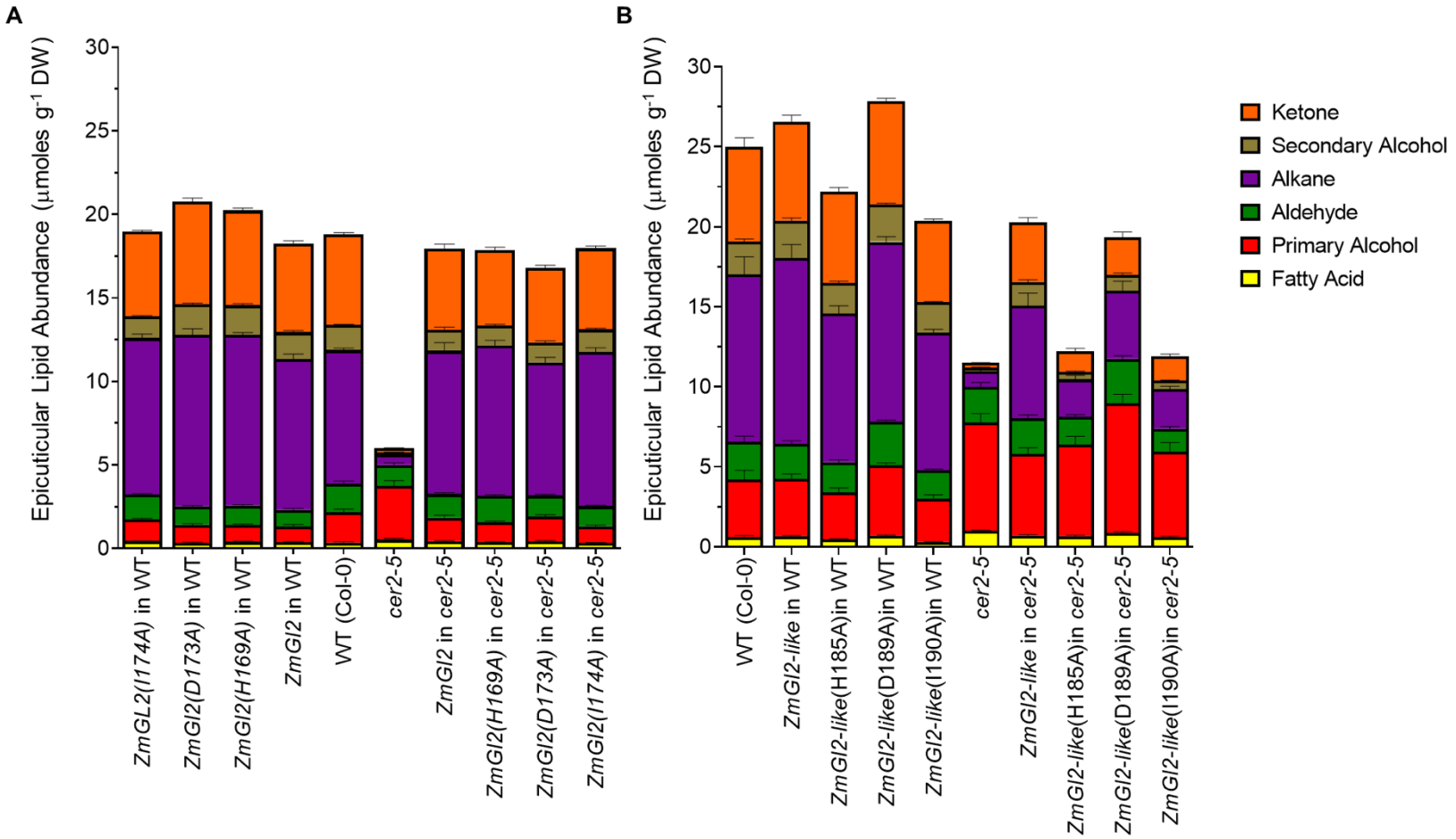
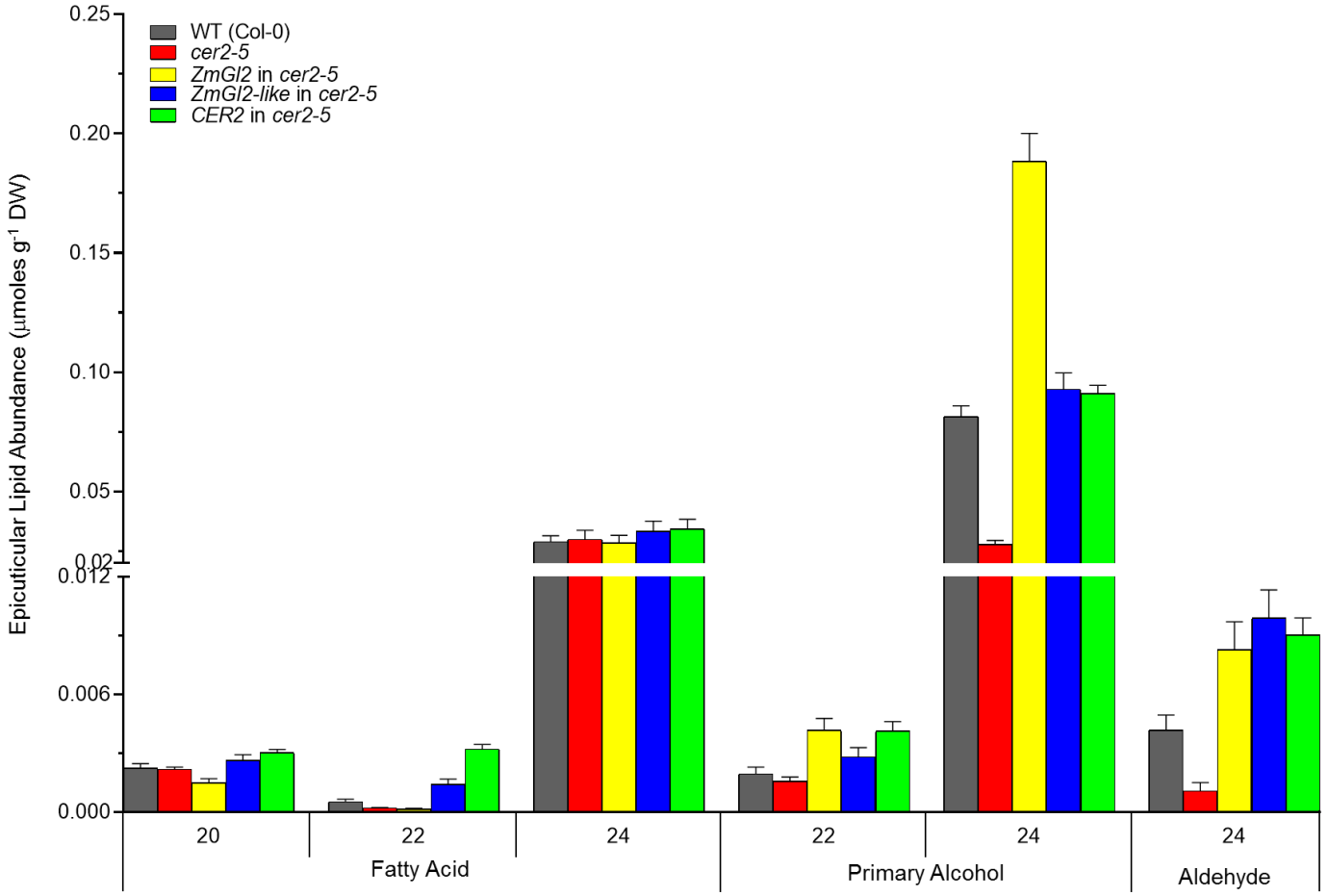
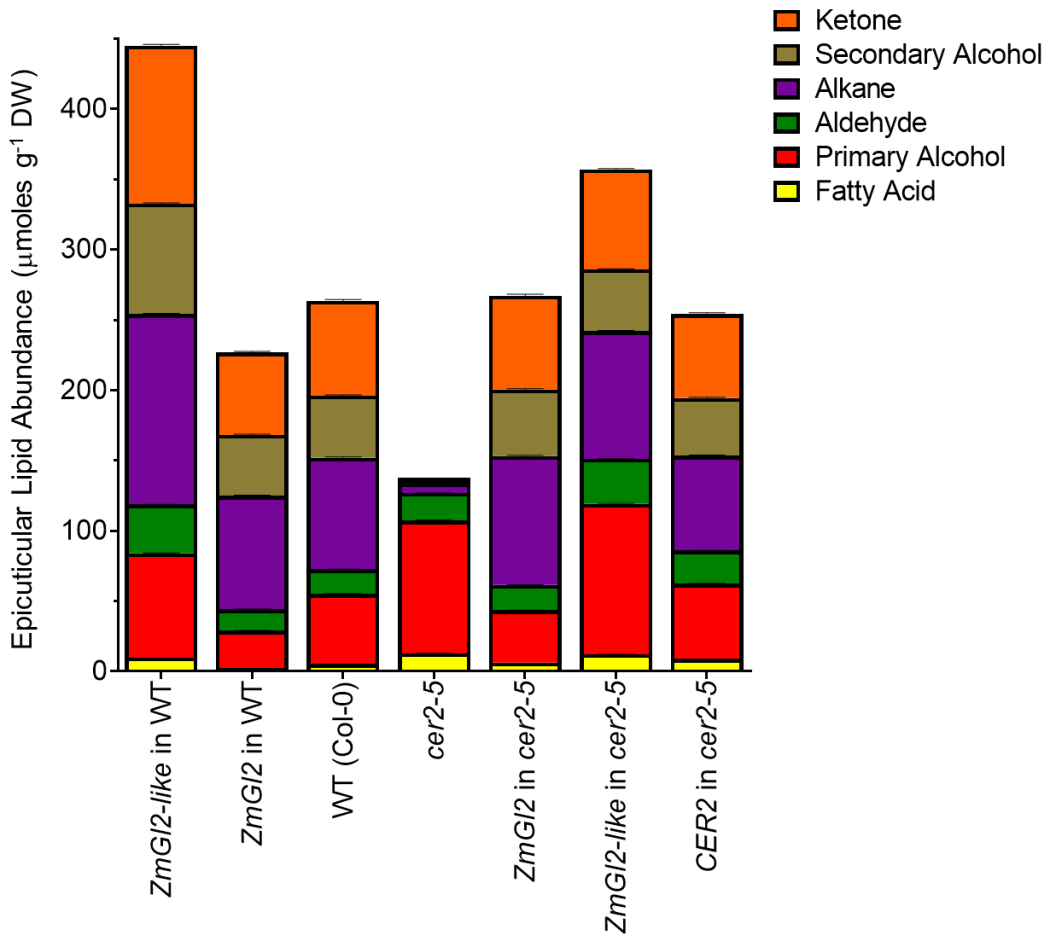


Figure 5. The role of the BAHD enzyme-defining, HXXXDX catalytic motif in supporting *in planta* functionality of *Gl2* (A) and *Gl2-like* (B) genes. Transgenes were either the wild-type *Gl2* or *Gl2-like*, or the indicated point mutants in the HXXXDX catalytic motif, and they were expressed either in a wild-type or *cer2-5* mutant genetic background. Epicuticular lipid compositions are compared among the identified genotypes, and each panel represents data collected from plant materials grown simultaneously in parallel. Epicuticular surface lipids were extracted from stems and analyzed as described in the Materials and Methods. Lipids are identified by their alkyl chemical class. All numeric data can be found in Supplemental Table 2. The data represent mean ± standard error of 10-40 replicates.

Supplemental Figures

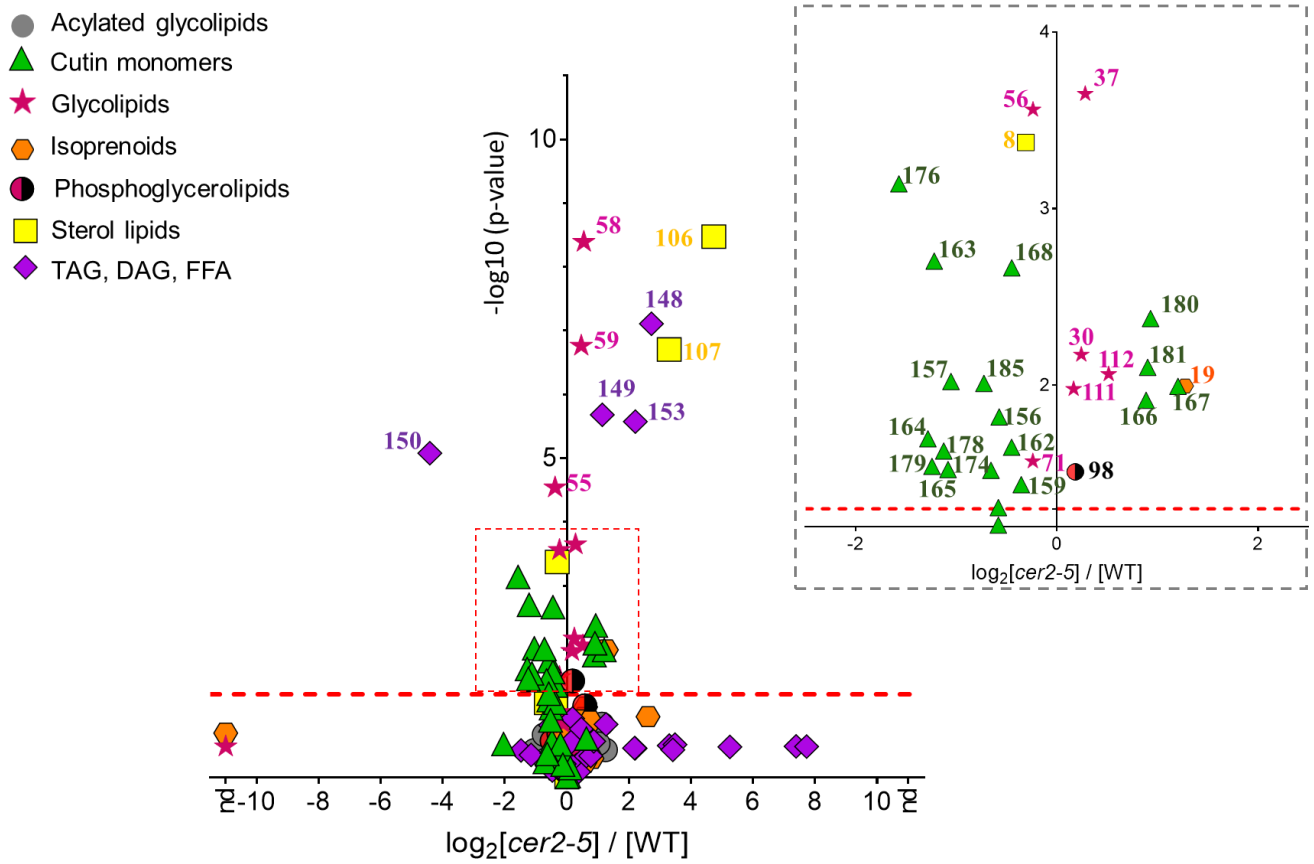


Supplemental Figure 1. Effect of transgenic expression of *Gl2*, *Gl2-like* and *CER2* on the extracellular epicuticular surface lipid profiles of stems of the *cer2-5* mutant plants (minor constituents of experiments presented in Figure 4A). Profiles are compared to the non-transgenic wild-type (Col-0) and *cer2-5* controls. Epicuticular surface lipids were extracted from stems and analyzed as described in the Materials and Methods. Lipids are identified by the carbon chain-length, and by the alkyl chemical class. Data for the major metabolites of C26-C32 carbon chain-length are in Figure 4A. All numeric data can be found in Supplemental Table 1. The data represent mean ± standard error of 10-15 replicates.



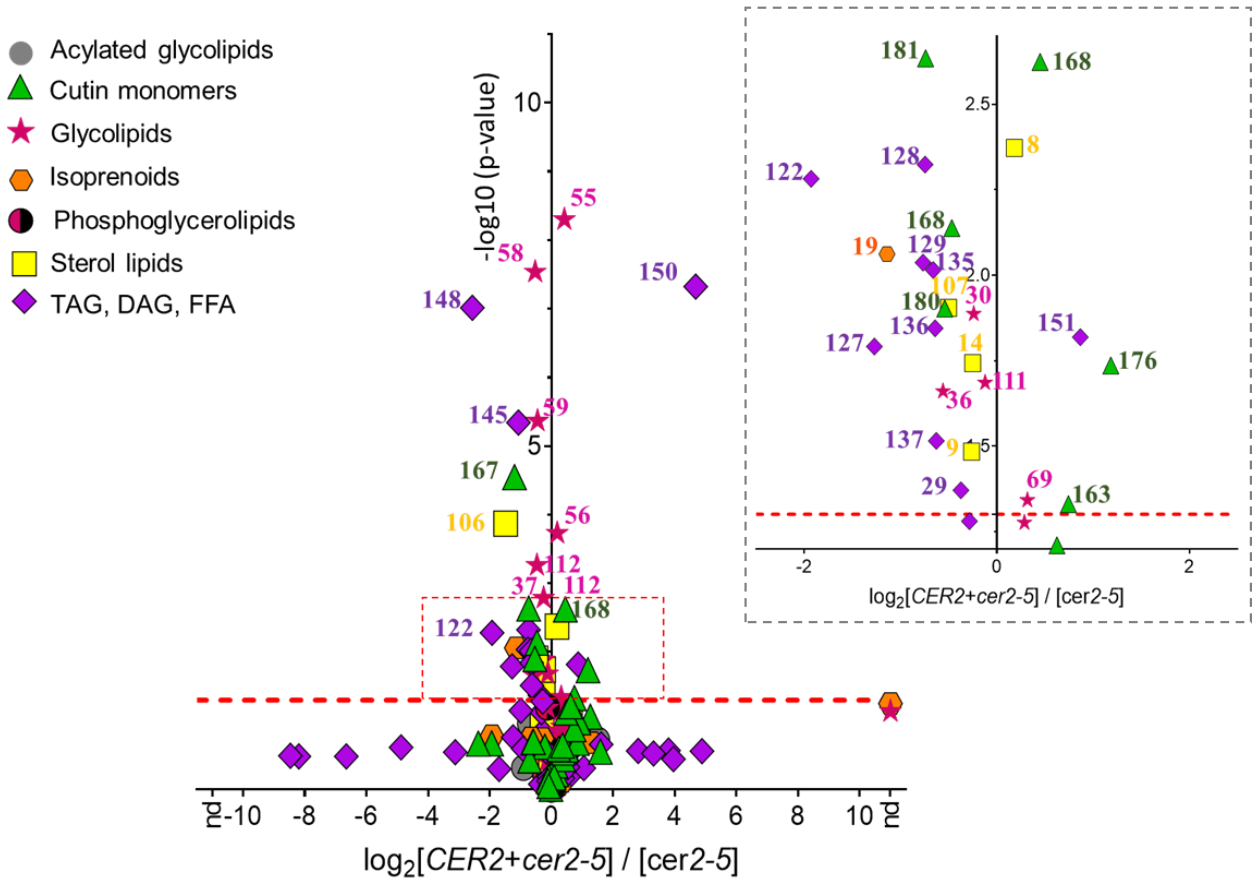
Supplemental Figure 2. Extracellular cuticular lipid composition of Arabidopsis stems. These data were obtained from the indicated genotypes, which includes the non-transgenic wild-type (Col-0) and *cer2-5* mutant, and these two lines transgenically expressing the *Glossy2* (*ZmGl2*), *Glossy2-like* (*ZmGl2-like*) or the *CER2* transgene. Data represents the average ± standard error (n= 10-15). The sums include alkyl chain lengths between C12 to C32.

(A)



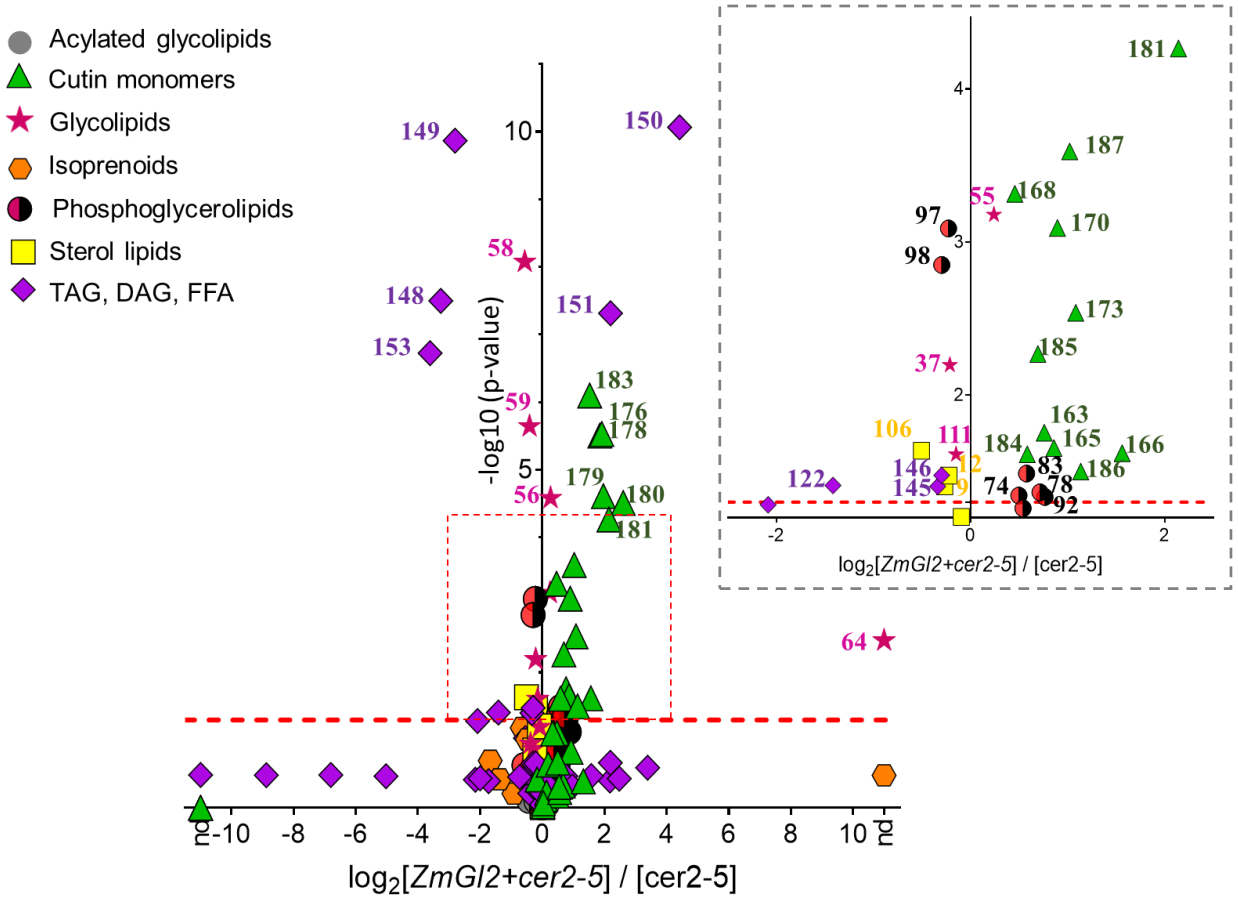
Supplemental Figure 3. Effect of transgenic expression of *Gl2*, *Gl2-like or CER2* on stem lipidome. The data are presented as Volcano-plots, with x-axis representing the log-ratios (base 2) of the relative abundance of individual metabolites as affected by a transgenic event, and the y-axis represents the statistical measure of significance (p-value), evaluated by Student's t-test; the data-points above the red dashed line are deemed statistically significant , with a p-value <0.05. The lipid class of each metabolite is identified by different data-symbols, and the numeral next to each data-point references the specific metabolite as identified in the numeric data presented in Supplemental Table 3. "nd" on x-axis indicates metabolites whose abundances are indistinguishable between the two genotypes. The transgenic events evaluated in each panel is identified in the label of the x-axis, and is as follows: (A) Non-transgenic plants comparing *cer2-5* mutant vs wild-type (Col-0), (B) Transgenic expression of *CER2* in the *cer2-5* mutant background vs *cer2-5* mutant, (C) Transgenic expression of *ZmGlossy2* in *cer2-5* mutant vs *cer2-5* mutant, (D) Transgenic expression of *ZmGlossy2* in wild-type vs wild-type (Col-0), (E) Transgenic expression of *ZmGlossy2-like* in *cer2-5* mutant vs *cer2-5* mutant, and (F) Transgenic expression of *ZmGlossy2-like* in wild-type vs wild-type (Col-0)

(B)



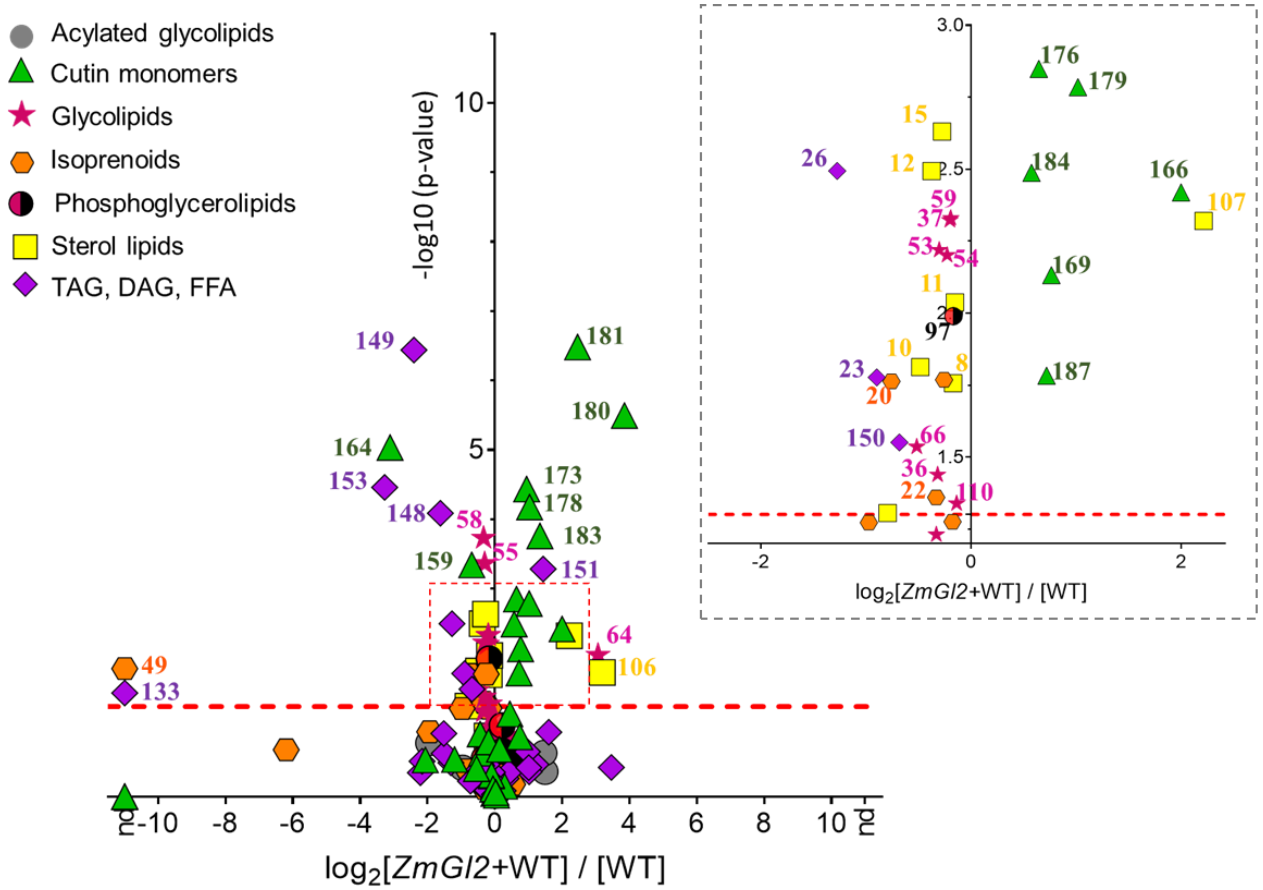
Supplemental Figure 3 Cont'd

(C)



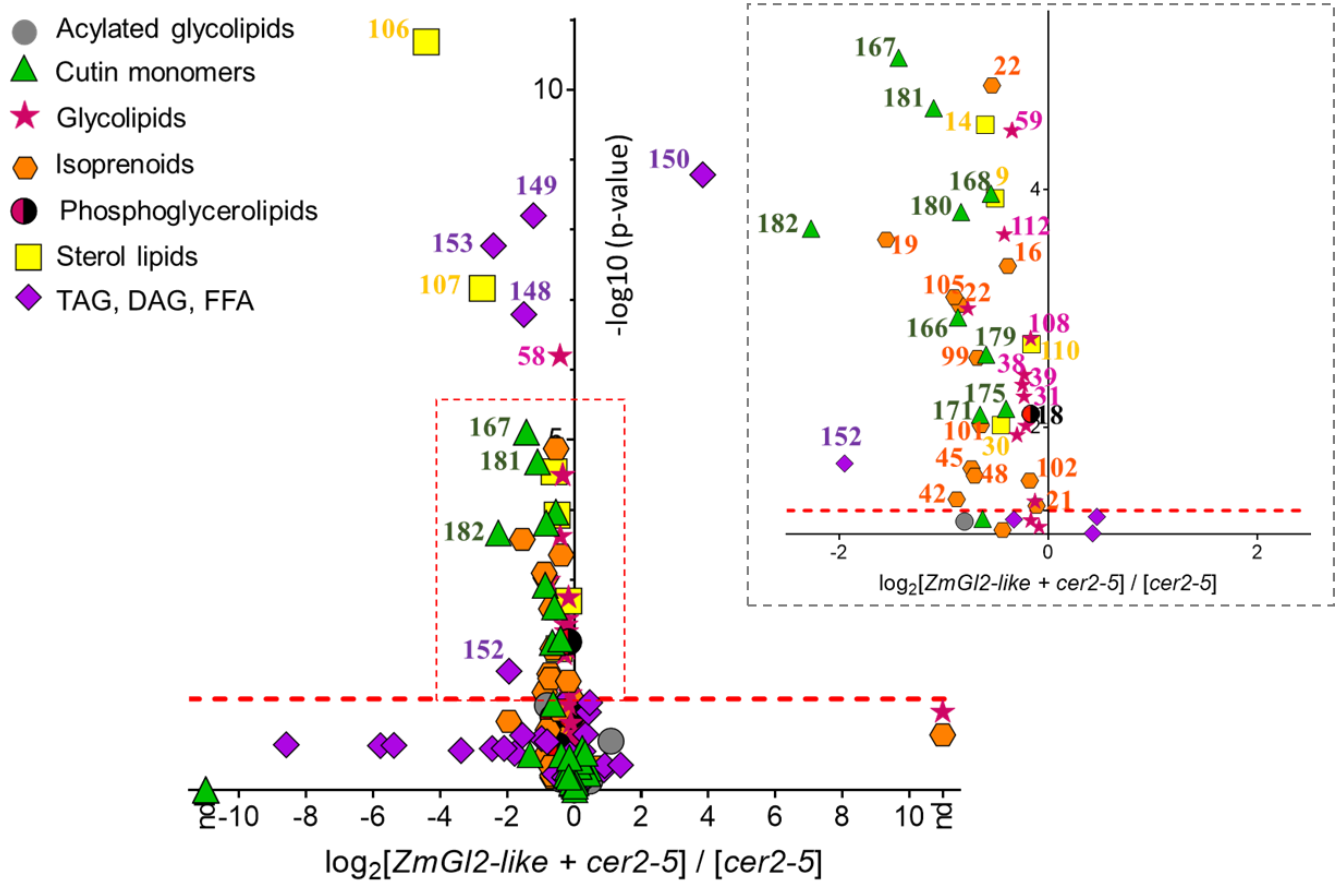
Supplemental Figure 3 Cont'd

(D)



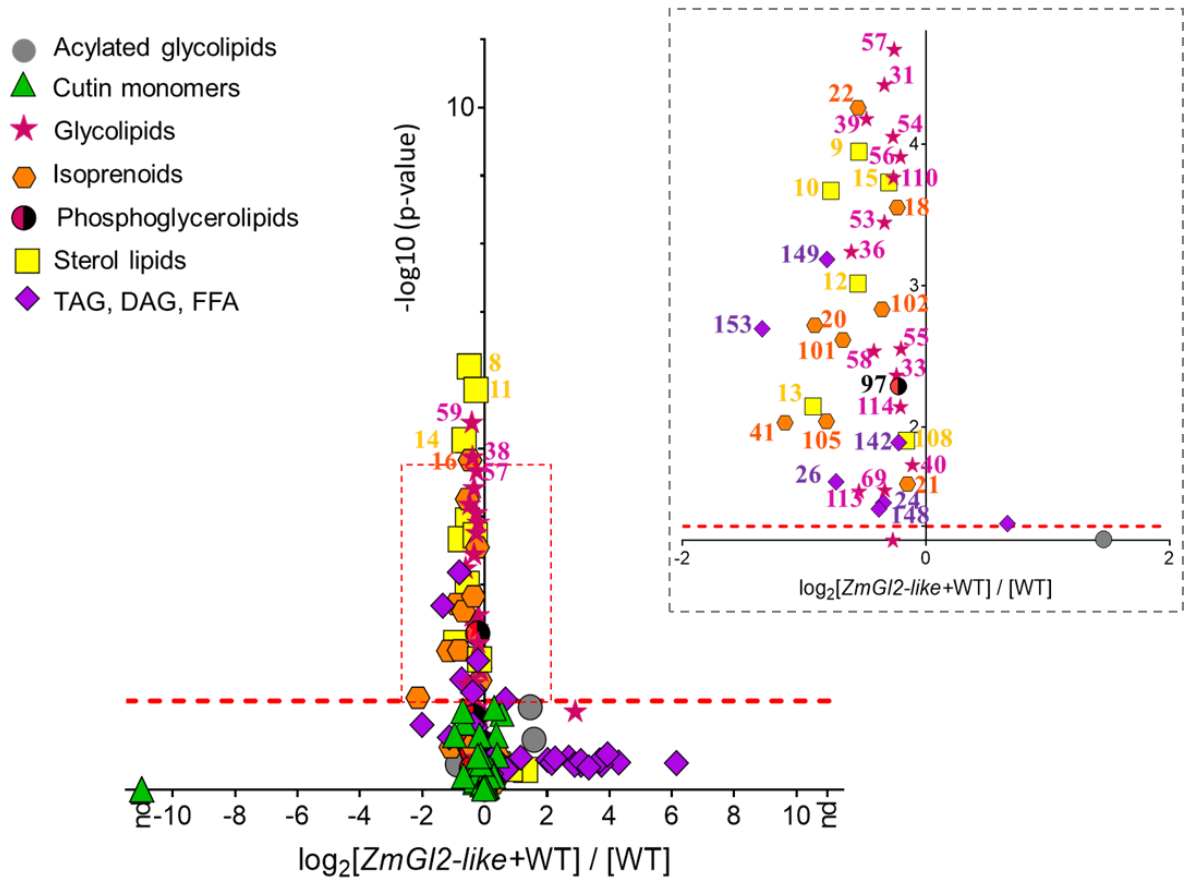
Supplemental Figure 3 Cont'd

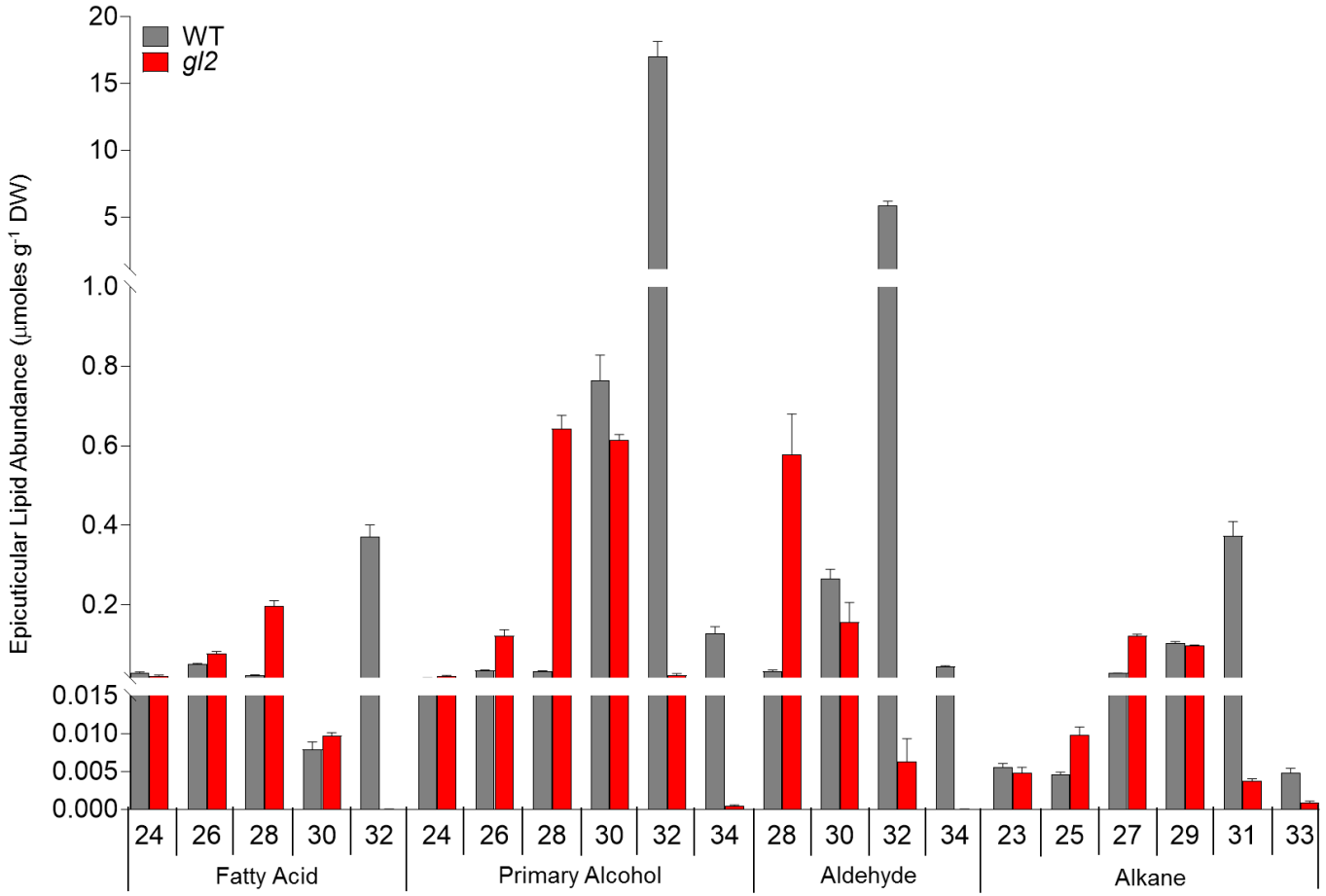
(E)



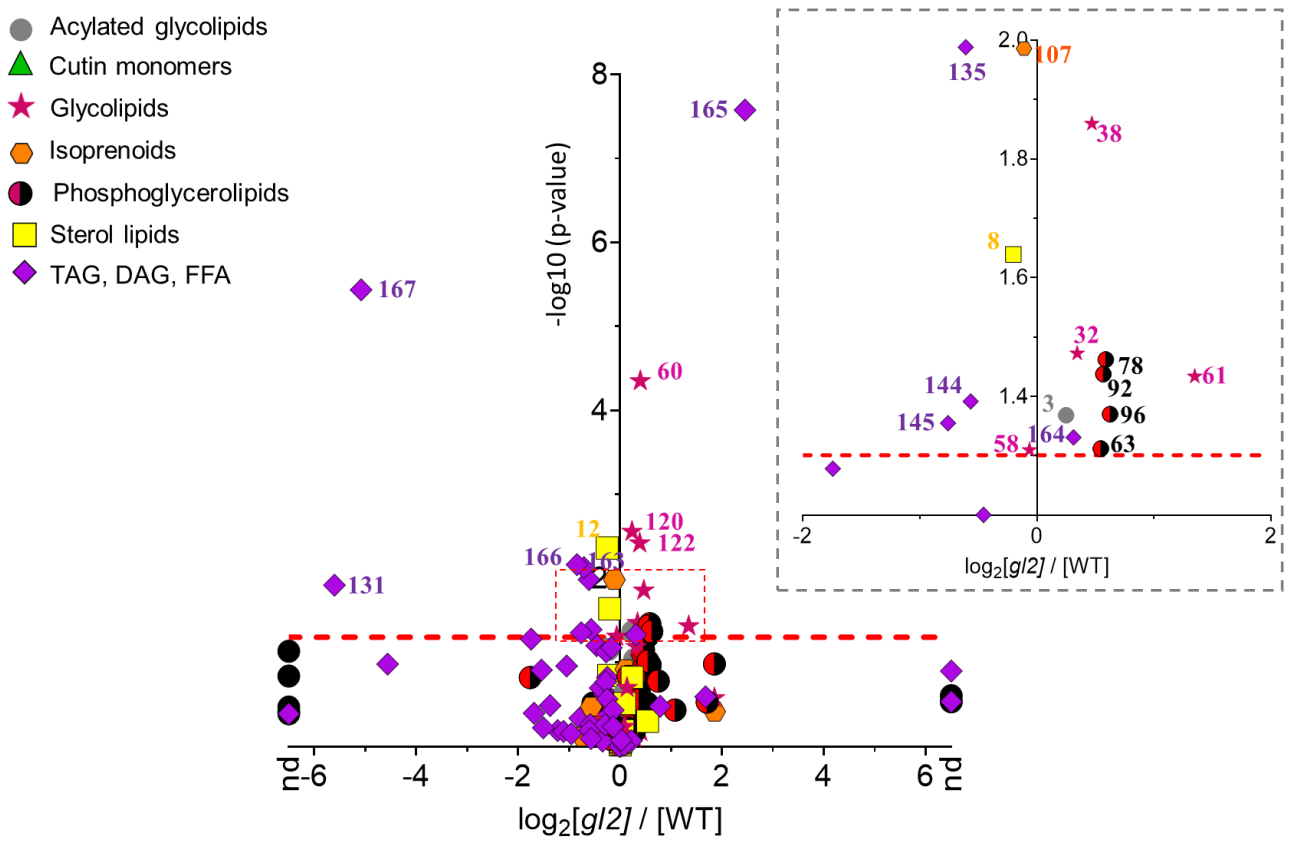
Supplemental Figure 3 Cont'd

(F)


Acylated glycolipids
Cutin monomers
Glycolipids
Isoprenoids
Phosphoglycerolipids
Sterol lipids
TAG, DAG, FFA
-log10 (p-value)
log2[ZmGl2-like+WT] / [WT]
nd
nd



Supplemental Figure 4. Effect of the *gl2* mutation on the on the extracellular epicuticular surface lipid profiles of seedling leaf (3-5 leaf stage) of maize. Epicuticular surface lipids were extracted from stems and analyzed as described in the Materials and Methods. Lipids are identified by the carbon chain-length, and by the alkyl chemical class. All numeric data can be found in Supplemental Table 1. The data were obtained from 4 pooled replicates each containing 5 different plants.



Supplemental Figure 5. Effect of the *gl2* mutation on the seedling leaf (3-5 leaf stage) lipidome of maize. The data are presented as Volcano-plots, with the x-axis representing the log-ratios (base 2) of the relative abundance of individual metabolites as affected by the *gl2* mutation, and the y-axis represents the statistical measure of significance (p-value), evaluated by Student's t-test; the data-points above the red dashed line are deemed statistically significant, with a p-value <0.05. The data were obtained from 4 pooled replicates each containing 5 different plants. The lipid class of each metabolite is identified by different data-symbols, and the numeral next to each data-point references the specific metabolite as identified in the numeric data presented in Supplemental Table 3. "nd" on x-axis indicates metabolites whose abundances are indistinguishable between the two genotypes.

CHAPTER 6. HIGH SPATIAL-RESOLUTION DISTRIBUTION OF EPICUTICULAR LIPIDS DURING THE DYNAMICS OF ARABIDOPSIS FLOWER DEVELOPMENT

Manuscript in preparation to be submitted to Plant Physiology

[1]Liza Esther Alexander[a,b], Jena S. Gilbertson[a], Bo Xie[c], Zhihong Song [b], Basil J. Nikolau[a,b].

Abstract

Collectively maize and Arabidopsis offer access to approximately 60 genetically defined genes that are involved in epicuticular lipid deposition, and these are known as the *glossy* or *eceriferum* loci, respectively. Although some of these loci have been molecularly defined, the specific functionality of these gene products is still unknown. In this study a combined transgenic and biochemical strategy was implemented to explore the physiological function of three homologous genes, the *Glossy2* and *Glossy2-like* genes of maize and the *eceriferum2* (*CER2*) gene of Arabidopsis. These biochemical characterizations integrated new high-spatial resolution mass spectrometric methods to profile the epicuticular lipids on developing floral tissues expressing either the *Glossy2* and *Glossy2-like* genes in wild-type or *cer2* mutant of Arabidopsis. Specifically, these methods included a) GC-MS analysis with large-volume injection using the programmable temperature vaporization technique enabling the analysis of epicuticular lipids extracted from individual flowers at discreet stages of development and b) matrix assisted laser desorption ionization mass spectrometric imaging that identified epicuticular lipid distribution

[a] Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA, USA.
[b] Center for Metabolic Biology, Iowa State University, Ames, IA, USA.
[c] Ames Laboratory-United States Department of Energy, Ames, IA, USA.

among floral tissues. Collectively, these datasets establish that both the maize *Glossy2* and *Glossy2-like* genes are functional homologs of the Arabidopsis *CER2* gene. Moreover, the 2 maize genes that were probably generated by an ancient gene-duplication event have neofunctionalized so that *Glossy2* has acquired additional functionalities in comparison to either *Glossy2-like* or *CER2*.

Introduction

The multicellular nature of plants has provided challenges to deciphering the biochemical mechanisms that regulate extracellular epicuticular lipid biogenesis. These lipids are produced by the epidermal cell layer that accounts for only about 10% of the cellular population of the aerial organs of plants (Jellings and Leech, 1982). Additionally, this biological process is divided among different compartments at the cellular and subcellular levels (cytoplasm, endoplasmic reticulum, and plasma membrane of epidermal cells). These technical barriers have been partially overcome with forward genetic approaches that use *eceriferum* (*cer*) mutants of Arabidopsis (Koornneef et al., 1989), *glossy* (*gl*) mutants of maize (Schnable et al., 1994) and tomato (Leide et al., 2007; Vogg et al., 2004), and *wax crystal-sparse leaf* (*wsl*) mutants of rice (Wang et al., 2017; Yu et al., 2008), which has led to the molecular identification and characterization of genes involved in extracellular lipid accumulation. Many examples indicate however, that the isolation of the causative gene that gives rise to the defect in epicuticular lipid deposition is only the starting point for detailed understanding of the function of the isolated gene (Bernard and Joubès, 2013; Post-Beittenmiller, 1996; Samuels et al., 2008).

These technical barriers can be partially overcome through advances in the fields associated with proteomics and metabolomics. A particular limitation with these fields is the fact that these techniques are generally performed on tissue extracts that homogenize tissues/cells that

are at different developmental stages, and thereby losing *in situ* localization and developmental states of the cell population. Although such technologies as reporter interactions (e.g., nucleic acid hybridization and antibodies) (Griffin et al., 1998; Küpper et al., 2007; McFadden, 1995), or gene and reporter molecule interactions (e.g., promoter-reporter or protein fusions to GFP and/or luciferase) (Chalfie et al., 1994; Gallagher, 1992; Koo et al., 2007; Ow et al., 1986) can detect high spatial and temporal scale resolution of gene expression, the products measured by these -omics platforms are not necessarily dynamic. This gap in the biochemical characterization pertaining to the nature and regulation of metabolism as intermediates of such processes move through metabolic networks in and among different cells, such limitations can be partially overcome by the use of more sophisticated technologies, such as mass spectrometry imaging (MSI) (McDonnell and Heeren, 2007).

MSI is a rapidly growing technology known for its ability to provide high spatial resolution data (~1μm spatial resolution), detecting extremely small amounts of metabolites (sensitivity), and providing chemical identification data concerning the detected analyte. Additionally, the use of diverse matrices as in the case of MALDI-MSI (matrix assisted laser desorption ionization) further enhanced the abilities of this technology, making it possible to analyze the spatial distribution of a variety of compounds, such as lipids, proteins, and small molecules in both plant and animal tissues (Amstalden van Hove et al., 2011; Angel and Caprioli, 2013; Jungmann and Heeren, 2012; Lee et al., 2012; Svatoš, 2010).

The specific example that is the focus of this study are the homologous maize *Glossy2* and *Glossy2-like* genes (Alexander et al., 2019; this thesis), whose specific biochemical functions are still unclear. In this study we illustrate the application of two specific analytical capabilities that attempt to overcome limitations associated with past characterizations, which are limited to

biochemical assays that measure metabolic outcomes in plant tissues taken at a single time-point, and thus ignore the dynamics of the system. Furthermore, the use of more assiduous analytical technologies enabled the detection and measurement of epicuticular lipid accumulation patterns in discreet tissue samples.

Specifically, we used an on-column large volume injection capabilities of the LVI-PTV injector (Engewald et al., 1999; Wilson et al., 2000) that concentrates dilute samples during the GC-injection process, and thus assayed the dynamics of epicuticular lipid accumulation during flower development. We coupled this capability with the use of matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-MSI) (Cha et al., 2009; Feenstra et al., 2017; Jun et al., 2010; Lee et al., 2012), to examine the changes in the distribution of epicuticular lipids on developing flowers, and obtained dynamic data at high spatial resolution scales.

Results

Developing flowers as an informative system for gaining insights on the dynamics of epicuticular lipid deposition

The relatively rapid rate of tissue differentiation and development that occurs during the emergence of flowers provides a convenient system to explore the dynamics of epicuticular lipid accumulation. In addition, using the LVI-PTV injector capabilities (Agilent Technologies; Engewald et al., 1999; Wilson et al., 1997) to concentrate dilute samples during the GC-injection process, we characterized the epicuticular lipid profiles extracted from individual flowers, collected at different stages of development. These attributes enabled the dynamic comparison of the epicuticular surface lipid accumulation of wild-type and *cer2-5* plants during flower development, and the effect of genetically complementing this mutation by the transgenic expression of the wild-type *CER2* allele.

The developmental stages of the flowers were defined from their morphological appearances, as follows: Stage A are flowers with closed buds; Stage B flowers are identified by the first emergence of petals from the bud; Stage C, the emerging petals are perpendicular to the flower axis; and at Stage D the flower recloses and the developing silique begins to emerge (Figure 1). Time lapse videos of 2 to 4 replicate flowers enabled the timing of these 4 stages, relative to Stage A, which was assigned the zero-time point. Thus, flowers on wild-type plants reached Stage B at 2.10 $\pm$ 0.05 h, Stage C occurred at 9.50 $\pm$ 0.96 h, and Stage D occurred at 33.50 $\pm$ 2.6 h. Statistical analysis of these data show that the B to C transition is slowed by the *cer2-5* mutation, but this delay is alleviated by the transgenic expression of *CER2* in this mutant (Figure 2; Supplemental Table1). In parallel, we determined the biomass of these flowers as a quantitative measure of flower growth and development (Figure 3; Supplemental Table 2). Thus, although the *cer2-5* mutation affects the developmental timing, it does not alter the growth of the flowers (Figure 3).

Analysis of the epicuticular lipid load on these developing flowers reveals dynamic changes in the accumulation of the alkyl metabolites. These metabolites are fatty acids and alkyl derivatives, such as primary alcohols, alkanes (linear and iso-branched), ketones, and secondary alcohols (Supplemental Data Sheet). The accumulation of these lipids increase asymptotically with development, through an initial rapid accumulation (i.e., the transition from Stages A to B) that is as high as 26.3 nmol of lipid/h/g of fresh weight, which transitions to a rate of 1.5-2.6 nmol of lipid/h/g of fresh weight between Stages C to D. The effect of the *cer2-5* mutation is to nearly eliminate the initial burst in epicuticular lipid accumulation and to reduce the later accumulation rate to near zero (Supplemental Figures 1 and 2; Supplemental Tables 3 and 4). As would be expected, the transgenic expression of *CER2* in the *cer2-5* mutant increases the rate of floral

epicuticular lipid accumulation to that of the wild-type state. In the context of these genetic-based changes in the dynamics of the total epicuticular lipid loads on the flowers, the major components are always hydrocarbons, which account for 70 to 80% of these lipids (Figures 4-6; Supplemental Data Sheet). In the wild-type state, over 70-80% of the hydrocarbons are primarily of 29-carbon chain length, and lesser amounts of linear chain or 2-methyl-branched chain hydrocarbons of 26, 28, 29, or 30 carbon atoms also occur.

The other less abundant classes of epicuticular lipids, in descending order of abundance, are secondary alcohols, ketones, primary alcohols and fatty acids. The accumulation patterns of the majority of the individual metabolites follow a similar asymptotic curve as the total epicuticular lipid accumulation pattern (Figures 4-6). Moreover, the chain length distribution of each of these alkyl derivatives match those of the hydrocarbons, as would be predicted from the currently accepted model of epicuticular lipid biosynthesis pathway, i.e., VLCFA elongation coupled to either a reductive or decarboxylative branched pathway (Bianchi et al., 1985; Post-Beittenmiller, 1996). The one exception is the chain-length distribution of the VLCFAs, which are dominated by 20 to 24 carbon chain length, whereas the other alkyl classes are dominated by the longer chain lengths (28-30 carbons).

Supplemental Tables 3 and 4 tabulates the rates of accumulation for each of the epicuticular lipid metabolites between each of the 4 developmental transitions. While the asymptotic rates of accumulation are maintained for most individual components in the *cer2-5* mutation, in the initial stages of development (Stage A-B transition) these rates are approximately ½ the rate determined in the wild-type plants. Two exceptions to this generalization are the net accumulation rates of C26 and C28 fatty acids and primary alcohols. The initial rate of accumulation of these metabolites

is increased in the *cer2-5* mutant by 2- to 5-fold as compared to the wild type, and this increase is reversed by the transgenic expression of *CER2* (Figure 4; Supplemental Table 4).

Transgenic expression of *Glossy2* and *Glossy2-like* transgenes restores the rates of epicuticular lipid accumulation on *cer2* mutant flowers

Using the above data as a baseline, two independent experiments were conducted to characterize the effect of transgenically expressing either the maize *Glossy2* or *Glossy2-like* genes. In each of these experiments the individual maize genes were expressed in both the wild-type background and in the *cer2-5* mutant background. The transgenic expression of both *Gl2* and *Gl2-like* in the *cer2-5* mutant does not affect the developmental timing of the flowers, however in the wild-type background they both induce a faster rate of flower development (Figure 2 and Supplemental Table 1). The ineffectiveness of *Gl2-like* to impact the timing of floral development, extrapolated to biomass gain of the flowers (Figure 3, and Supplemental Table 2). In contrast, the transgenic expression of *Gl2* in either the wild-type or *cer2-5* mutant backgrounds suppresses the growth of the flowers at different developmental stages, namely the transition between Stages A to B, and Stages C to D, respectively (Figure 3, and Supplemental Table 2).

The effect of the transgenic expression of either *Gl2* or *Gl2-like* in the *cer2-5* mutant is to increase the level of epicuticular lipid accumulation on developing flowers. The *Gl2* transgene appears to be more effective in returning the accumulation of epicuticular lipid to wild-type levels (Supplemental Figure 1), whereas the *Gl2-like* transgene only partially rescues this trait (Supplemental Figure 2). This latter effect is particularly noticeable in the accumulation pattern of the major components, the alkanes, ketones and secondary alcohols (Figure 6). Thus, whereas the transgenic expression of *Gl2* in the *cer2-5* mutant restores the levels of these lipids to wild-

type levels, the transgenic expression of *Gl2-like* in this mutant leads to the accumulation of these lipids at only 50% of the wild-type levels (Figures 5 and 6).

Mass spectral imaging reveals a non-uniform tissue and cellular level spatial distribution of epicuticular surface lipids on the abaxial and adaxial surfaces of Arabidopsis flowers

The dynamic data presented previously (Figures 4-6 and Supplemental Figures 1 and 2), integrates the epicuticular lipid compositional changes that occur on a number of different flower tissues (e.g., petals, sepals, anthers and stigmatic tissues), each of which is on a different developmental schedule (Alvarez-buylla et al., 2010; Smyth, 1990). Therefore, we applied *in situ* mass spectrometric imaging, specifically MALDI-MSI technology to individually localize these metabolites at a high spatial scale (~80 µm) and thus enable the visualization of the accumulation of these metabolites at the level of different tissues and possibly a few cells (2-4 cells). Moreover, MALDI-MSI was used to investigate the epicuticular surface lipid distribution on both adaxial- and abaxial surfaces of the developing flowers (Figure 7A). The images show the metabolite distribution of individual metabolites on the adaxial surfaces of sepals and petals, the outer surface of anthers, and the stigmatic surface, and MSI of the abaxial surface imaged only the outer facing surfaces of the petals and sepals of each flower (Figure 7B).

MSI technology was used to compare all four developmental stages of wild-type Arabidopsis flowers, and these were compared to flowers of *cer2-5* mutant and to the *cer2-5* mutant that are transgenically expressing the *Gl2* gene (Figure 7B). These data reveal that the most abundant epicuticular lipid (C29 alkane) is evenly distributed on both the adaxial and abaxial floral surfaces of wild-type flowers during the later stages of development (Stages C and D). However, at the earlier stages of foliar development (Stages A and B), this epicuticular lipid is only detectable on the abaxial surface and is reduced on the adaxial surface of the flower (Figure

7B). The *cer2-5* mutation significantly reduced the amount of C29 alkane on both the adaxial and abaxial floral surfaces and the overexpression of *Gl2* in this mutant background reversed this effect, so that the distribution is like the wild-type.

Because Stage C flowers display the clearest separation of the different flower tissues, more detailed MSI analysis was used to evaluate the effect of transgenically expressing *Gl2* and *Gl2-like* in the *cer2-5* mutant background (Figure 8). The spatial distribution of C30 fatty acid and its derivatives (C30 aldehyde, and C29 alkane, secondary alcohol and ketone), demonstrate that whereas these metabolites are evenly distributed on the abaxial surface, they more specifically localize to the petal, and edges of the sepals of the adaxial surface of wild-type flowers. The *cer2-5* mutant flowers show even, but reduced levels of C30 aldehyde, C29 alkane and C29 ketone on both surfaces of the flower sepals and petals, and the C30 fatty acid and the C29 secondary alcohol are below detection levels in the mutant. The transgenic expression of *Gl2* or *Gl2-like* in the *cer2-5* mutant not only increased the levels of these metabolites, but also reversed the distribution pattern of all these metabolites to the wild-type patterns (Figures 8A and 8B).

Use of MSI data to visualize the spatial distribution of potential individual metabolic reactions of the epicuticular biosynthetic pathway

An additional application of MSI is the ability to infer the underlying metabolism that supports the distribution of metabolites at a high spatial resolution. We illustrate this by the visualization of the co-localization of substrate-product pairs of individual metabolic reactions of the epicuticular biosynthetic pathway. For example, consider the distribution of metabolically-related metabolites of fatty acid elongation from C24-C32 and their alkyl-derivatives. Figure 9 and Supplemental Figure 3 and 4 illustrate the visualization of these relationships by false color coding the substrate (green) and product (red) of a metabolic reaction, and the resulting fusion of these

two images. In such fused images, yellow colored zones indicate the underlying tissue/cell locations where this metabolic interconversion could be occurring.

Thus, Figure 9, Panel A show that C24 and C26, C26 and C28, and C28 and C30 fatty acid pairs (precursor-product pairs of fatty acid elongation reactions) predominately co-localize on the adaxial surface of wild-type flower petals. In contrast, there is less such co-localization on the abaxial surface (panel B Figure 9). Similar co-localization occurs on the surface of the stigmatic tissue, petals, and abaxial surface of the sepals for the substrate-product pairs for the three reactions that lead from the C30 fatty acid to the C29 ketone (Figure 9B). One can contrast this latter finding with the co-localization of substrate-product pairs of the reductive pathway that generates the primary alcohols. Namely, the C28 aldehyde-C28 primary alcohol and C30 aldehyde-C30 primary alcohol that co-localize on the stigmatic tissue (Figure 9A).

The *cer2-5* mutation does not affect these co-localization patterns, with the exception of products that are generated from C28 and C30 fatty acids. Thus, consistent with the proposed role of CER2 in affecting the C28 to C30 fatty acid conversion (Hannoufa et al., 1993; Jenks et al., 1995), the C28-C30 fatty acid substrate-product pairs do not co-localize in the *cer2-5* mutant (Supplemental Figure 3A). As a result, there is altered accumulation patterns of the C28 alkyl derivatives within both the decarboxylative and reductive branches of the epicuticular lipid pathway. Specifically, the substrate-product pairs from C28 aldehyde to C27 ketone primarily co-localize in the sepals, and the C26 aldehyde-C26 primary alcohol and C28 aldehyde-C28 primary alcohol product-substrate pairs primarily co-localize in the sepals and stigmatic tissue surface. These differential colocalizations may indicate that both the decarboxylative and reductive branches of the epicuticular pathway occur on sepal tissues, but the reductive pathway appears to predominate in the stigmatic tissues. The transgenic expression of *Gl2-like* in the *cer2-5* mutant

background, restores the co-localization pattern of the C28-C30 fatty acid elongation product-substrate pairs and the decarboxylative or reductive pathway product-substrate pairs to the pattern observed in the wild-type flowers (Supplemental Figure 3B).

These substrate-product pair co-localization patterns were also determined and compared at different stages of flower development (Supplemental Figure 4). These data are presented for the abaxial surfaces of the flowers because the signal was too low to reliably gather such data for the adaxial surface. In wild-type flowers, during the development from Stages A to D, the co-localization patterns of the fatty acid elongation substrate-product pairs (C24 to C28 fatty acids) are reduced as the flowers proceed to Stage D, but colocalization of the C26 to C30 fatty acid substrate-product pairs co-localize to the sepals (Supplemental Figure 4 A-D). The *cer2-5* mutation disrupts these co-localization patterns, so that only the C24/C26 fatty acid pair colocalize in petals, whereas, the C26/C28 fatty acid pair co-localize to the sepals throughout flower development (Supplemental Figure 4 E-H). The transgenic expression of *Gl2* in the *cer2-5* mutant did not affect these altered co-localization patterns of the C24/C26 fatty acid substrate-product pairs, and no co-localization patterns were observed for C26/C28 fatty acid pair later in flower development (Stages C and D) (Supplemental Figure 4 I-L).

The substrate-product pairs for the decarboxylative and reductive branches of the pathway are distinct from the fatty acid elongation pairs; namely, they colocalize to the abaxial surfaces of sepals. This colocalization pattern is most apparent at early developmental stages of the flowers (Stages A and B; Supplemental Figure 4 A and B), but becomes weaker as the flowers develop through Stages C and D (Supplemental Figure 4 C and D). The *cer2-5* mutation alters these co-localization patterns, particularly for the C25 and C27 alkane to ketone substrate-product pairs, and the C26 aldehye-C26 primary alcohol substrate-product pairs. Specifically, the sepal co-

localization patterns for C25 alkane to ketone are maintained in the early stages of flower development (Stages A and B) (Supplemental Figure 4 E and F), whereas the C27 alkane to ketone pairs, and C26 aldehyde-primary alcohol pair are maintained throughout development (Supplemental Figure 4 E-H). These highlighted *cer2*-dependent alterations in the co-localization of substrate-product pairs are consistent with the proposed role of CER2 in affecting the C28 to C30 fatty acid conversion (Hannoufa et al., 1993; Jenks et al., 1995). The transgenic expression of *Gl2* in the *cer2-5* mutant background restored sepal co-localization patterns for the C30 alkyl decarboxylative and reductive pathway substrate-product pairs, at least at the early stages of flower development (Stages A and B) (Supplemental Figures 4 I and J) but not in the later stages of development (Stages C and D) (Supplemental Figures 4 K and L).

Discussion

Elucidating the biochemical mechanisms regulating metabolic processes are complicated by the multicellular nature of plants. Specifically, molecular machineries are distributed among different subcellular and/or cellular compartments that integrate to achieve net metabolic interconversions. An additional complexity associated with deciphering the epicuticular lipid biosynthetic pathway is associated with the fact that it's restricted to the epidermal cells of aerial plant organs, and the products of the pathway are unidirectionally transported to the outer surface of plant organs. Forward genetic approaches have proven useful in isolating and characterizing genes pertaining to epicuticular lipid biogenesis, including *glossy* (*gl*) loci of maize (Schnable et al., 1994). Characterization of *Glossy2* and the recently identified *Glossy2-like* gene has been accomplished by genetic complementation assays, relative to the function of the Arabidopsis *CER2* gene (Alexander et al., 2019; this thesis). As with other molecular genetic characterizations of *glossy* and *cer* genes, data are interpreted relative to a core of epicuticular lipid biosynthesis

pathway, which integrates fatty acid elongation followed by either a reductive branch or a decarboxylative branch, which appear to be common to both maize and Arabidopsis (Bianchi et al., 1978; Bianchi et al., 1985; Kolattukudy, 1970; Kunst and Samuels, 2009; Lee and Suh, 2013; Post-Beittenmiller, 1996).

Dynamic distribution of epicuticular surface lipids

The majority of past genetic-based studies of the accumulation of epicuticular lipids have primarily been limited to a single time-point of analysis, and overlook the dynamics of the processes that integrate the developmental program of the tissue/organs that are being evaluated. Furthermore, most such data are collected from the analysis of epicuticular lipid extracts prepared from relatively large biological samples that combine lipids generated by a large collection of cells and tissues that are most probably of different age and different developmental states. Emerging flowers and the application of new, more assiduous analytical technologies offer the potential to address both of these potential limitations in deciphering the physiology of epicuticular lipid accumulation patterns.

Specifically, we investigated the effect of the *cer2* mutation and the transgenic expression of *Gl2* and *Gl2-like* transgenes on epicuticular lipids of flowers, monitoring the accumulation pattern of the lipids during the course of floral development. These data were gathered with epicuticular lipid extracts prepared from individual flowers at four different stages of development. In parallel, we also applied MALDI-MSI to image the distribution of epicuticular lipid metabolites on either the abaxial or adaxial surfaces of the developing flowers at a sufficiently high spatial resolution to distinguish between four different floral tissues (petals, sepals, anthers and stigmatic surfaces) and determined the effect of each transgenic manipulation on the extracellular epicuticular lipid profiles.

The studies of the dynamics of flower development demonstrate that a block in a metabolic process that specifically occurs in the epidermal cells of the flower tissues (i.e., epicuticular lipid deposition caused by the *cer2* mutation), does not affect the developmental timing of the integrated organ, but appears to affect the growth of the organ, particularly when the *cer2* mutation is complemented by the maize homologs. Prior characterization of *cer2* mutants have been interpreted to indicate that this gene product affects the elongation of fatty acids between C26/C28 to C30 (Hannoufa et al., 1993; Jenks et al., 1995). The expression of the *Gl2-like* and *Gl2* transgenes in the *cer2-5* mutant reversed the effect on the accumulation of epicuticular lipids on flowers, but these transgenes have differential effects on the timing of flower development and/or growth of this organ. These findings indicate that the flower biomass trait can be affected by the deposition of the epicuticle.

During flower development the majority of the individual epicuticular lipid components accumulate at near parallel, asymptotic patterns as the flower developed. The maximum observed rate of epicuticular lipid accumulation occurs during the earliest stages of flower development, and this is at a rate of ~26 nmol/h/g fresh weight. This rate defines the minimal rate of *de novo* fatty acid biosynthesis that is required to feed the FAE system with C18-acyl-CoA precursor. Furthermore, because the final products of the epicuticular lipid biosynthetic pathway are components that are predominantly derivatives of 30-carbon fatty acids, 6 elongation cycles of the FAE system are required to elongate the C18 fatty acid precursor to these final products. Because each elongation cycle requires 1 molecule of malonyl-CoA generated from the cytosolic acetyl-CoA pool, and 2 molecules of NADPH, one can infer the minimal rates of acetyl-CoA generation in the cytosol (156 nmol/h/g fresh weight) that is required to support the observed rate of epicuticle lipid deposition, and similarly 312 nmol/h/g fresh weight of NADPH generation. Prior

characterizations of ATP-citrate lyase (Fatland et al., 2002; 2005), which generates the cytosolic acetyl-CoA pool, appears to indicate that there is sufficient activity of this enzyme in Arabidopsis tissue to provide this rate of acetyl-CoA. Similarly, there appears to be sufficient activities of NADPH-generating enzymes (NADP+-dependent malic enzyme, glucose-6-phosphate dehydrogenase, and 6-phosphogluconate dehydrogenase) to support these needs by FAE (Wakao et al., 2008; Wakao and Benning, 2005; Wheeler et al., 2005; Yin and Ashihara, 2008)

The effect of the *cer2-5* mutation is to reduce the initial rate of epicuticular lipid accumulation by about 50% (to ~14 nmol/h/g fresh weight), and this rate plateaus to near zero later in flower development. The transgenic expression of *CER2* in the *cer2-5* mutant background reverses the initial rate to that of wild-type state, enabling the normal levels of epicuticular lipid accumulation. Similarly, the transgenic expression of *Gl2* (but not *Gl2-like*) also reverses the effect of the *cer2* mutation, increasing the rate of epicuticular lipid accumulation to that of wild-type levels. In contrast, *Gl2-like* transgene reversed the effect to about half that of wild-type.

Furthermore, in these genetically complemented lines (i.e., the *cer2-5* lines expressing either *Gl2, Gl2-like* or *CER2* in the *cer2-5* mutant) the near asymptotic pattern of accumulation is shared by all individual epicuticular lipid components, and these are at higher rates than in the *cer2-5* mutant. The exception to this generalization is the pattern of accumulation displayed by the primary alcohols and fatty acids. Compared to the wild-type controls, the net accumulation rates of these latter metabolites are enhanced in the *cer2-5* mutant and this alteration is reversed only by the *Gl2-like* transgene. In combination, these genetic complementation experiments indicate that the 2 maize genes have overlapping and distinct functionalities, but *Gl2*-functionality is more homologous to the *CER2* gene function.

High spatial resolution distribution of epicuticular surface lipids

As visualized by MSI, the spatial distribution of the epicuticular lipid components acquire unique tissue localizations. Specifically, epicuticular lipid deposition is distinct between the adaxial and abaxial floral surfaces, with the latter surface having greater levels of deposition. The larger accumulation of epicuticular lipids on the abaxial side of the flower, primarily localized to the sepals, is consistent with the function of the cuticle as a water barrier (Aarts et al., 1995; Jung et al., 2006; Millar et al., 1999). Namely, because the sepals enclose the flower bud, their abaxial surfaces are outward facing, and therefore the larger epicuticular lipid load would provide more extensive protection from the desiccating environment

Flowers transgenically expressing the *Gl2* or *Gl2-like* transgenes proved useful in visualizing the spatial distribution of the epicuticular lipid biosynthetic reactions. This was accomplished by the co-localizing spectral images of substrate-product pairs of specific reactions of different segments of the epicuticular lipid pathway. Thus, MSI analysis of *Gl2,* and *Gl2-like* expressing flowers demonstrate that the core process of fatty acid elongation occurs predominantly in the petals of flowers. Subsequently, the products of the FAE system are primarily processed through the decarboxylative or reductive branch of the pathway. The co-localization spectral images of substrate-product pairs indicate that the decarboxylative branch is expressed in the stigmatic tissue and sepals, and the reductive branch of the pathway is primarily expressed in the stigmatic tissue.

Material and Methods

Plant Material and Growth Conditions

T-DNA mutant line SALK_084443C (*cer2-5*; At4g24510) in the Col-0 background, wild-type Col-0, and both these lines over-expressing maize transgenes *Glossy2* (GRMZM2G098239)

and *Glossy2-like* (GRMZM2G315767), along with *CER2* overexpressed in *cer2-5* mutant background, were generated and grown as previously described in CHAPTER 5 (Alexander et al., 2019; this thesis). Biochemical analyses were conducted on developing flowers of Arabidopsis plants of different genotypes harvested at four stages of development. Stage A are flowers with closed buds; Stage B flowers are identified by the first emergence of petals from the bud; Stage C, the emerging petals are perpendicular to the flower axis, and at Stage D the flower recloses and the developing silique begins to emerge. *Glossy2* and *Glossy2-like* transgenic expression analysis were conducted as independent experiments.

Epicuticular lipid analysis of single Arabidopsis flowers

Extracellular epicuticular lipids were extracted from individually collected flowers based on developmental growth phases (i.e, Stages A to D). Time-lapse photography was used to determine the timing of the transition between each stage of flower development and weight of single flowers at four stages of development were estimated from 2-3 biological replicates consisting of 6-13 pooled flower stages. For biological replication data were gathered from six individual flowers at each stage of development, each flower being harvested from independent plants. For quantification purposes, an aliquot of 0.01-µg hexacosane was applied to the individually harvested flowers, and they were immersed in 0.5-mL chloroform for 60s. The abundance of the extracellular lipids were calculated based on this internal standard. The extracellular cuticular lipid extracts were dried under a stream of N_2 gas and derivatized using 0.07-mL *N*, *O*-Bis (trimethylsilyl) trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS) (Sigma-Aldrich) and incubated at 70°C for 30min.

GC-MS analysis was performed with an Agilent 6890 GC interfaced to a 5973 mass spectrometer. The GC was equipped with a LVI-PTV inlet injector (Agilent Technologies), which

enabled the analysis of dilute samples by providing a means of concentrating large volume samples (10-µL) during the injection process. The GC was equipped with a HP-5ms column (30 m x 0.25 mm i.d. coated with a 0.25 µm film, Agilent Technologies), with a mobile gas-phase of helium at a flow rate of 0.1 mL/min. The column oven temperature was programmed to increase from 80°C to 180°C at 20°C/min, and held at this temperature for 1min, then ramped to 220°C a rate of at 5°C/min, held at this temperature for 5min, and finally ramped to 320°C at 10°C/min, and held there for 10min. The interface temperature to the mass-spectrometer was at 280°C, with an ionization voltage of 70 eV.

The GC/MS data files were deconvoluted by National Institute of Standard and Technology Automated Mass spectral Deconvoluted and Identification System (NIST AMDIS) software, and searched against an in-house compound library as well as the NIST 14 Mass Spectral Library.

Mass Spectrometry imaging (MSI) of flowers

Flowers collected at the four developmental stages were attached to glass slides (FisherScientific) with double-sided tape. Flowers were carefully opened for imaging the adaxial surfaces of the floral organs, whereas for the abaxial (lower) surface, petals and sepals were carefully dissected and placed on the double-sided tape. Care was taken to ensure that all flower organs were flat on the MS imaging slide. The attached samples are dried in a desiccator for 30-60 min. Flowers were then sputter-coated with silver target 3NS (99.95% purity; ESPI Metals, Ashland, OR) for 50s, using a Cressington 308R Sputter Coater (Redding, CA, USA) at an argon partial pressure of 0.02 mbar and a current of 80 mA.

A 7T SolariX Fourier Transform Mass Spectrometer (FTICR-MS) (BrukerDaltonics, Bremen, Germany) equipped with an MALDI ion source that used a SmartBeam^{TM} II laser with a spot size of 80 µm diameter was utilized for MS imaging. All spectra were acquired in positive ion

mode integrating 200 laser shots per imaging pixel. The laser power was set to 25% with a frequency of 1000 Hz. The mass scanning range was set from m/z 300 to m/z 800, time of flight value was 0.7 s, and ion cooling time was set to 0.01 s. The mass spectrometer was externally calibrated prior to sample analysis, using arginine standard mixture. Images were developed using FlexImaging 4.0 software (BrukerDaltonics, Bremen, Germany). Chemical identification of ions were performed based on theoretical monoisotopic masses of each surface lipid species and the mass tolerance was set to 2 ppm.

Acknowledgements

The authors acknowledge Drs. Ann Perera, and Lucas Showman of the WM Keck Metabolomics Research Laboratory (Iowa State University, Ames, IA) for assistance in metabolic and mass spectral imaging analyses; and Dr. Yue Wu (Iowa State University, Ames, IA) for the use of sputter coater equipment. This work was partly funded by US Department of Energy and by the State of Iowa through the Center for Metabolic Biology.

References

Aarts, M.G., Keijzer, C.J., Stiekema, W.J., Pereira, A., 1995. Molecular characterization of the CER1 gene of arabidopsis involved in epicuticular wax biosynthesis and pollen fertility. Plant Cell 7, 2115–2127.

Alvarez-buylla, E.R., Benítez, M., Corvera-poiré, A., Cador, C., Folter, S. De, Buen, A.G. De, Garay-arroyo, A., García-ponce, B., Jaimes-miranda, F., Rigoberto, V., Piñeyro-nelson, A., Sánchez-corrales, Y.E., 2010. Flower Development 1–57.

Bernard, A., Joubès, J., 2013. Arabidopsis cuticular waxes: Advances in synthesis, export and regulation. Prog. Lipid Res. 52, 110–129. https://doi.org/10.1016/j.plipres.2012.10.002

Bianchi, A. (1978). Glossy mutants: Level of action and level of analysis. Maize Breeding and Genetics, D.B. Walden, ed (New York: John Wiley and Sons), pp 533-550

Bianchi, A., Bianchi, G., Avato, P., Salamini, F. 1985. Biosynthetic pathways of epicuticular wax of maize as assessed by mutation, light, plant age and inhibitor studies. Maydica 30, 179-198.

Cha, S., Song, Z., Nikolau, B.J., Yeung, E.S., 2009. Direct Profiling and Imaging of Epicuticular Waxes on Arabidopsis thaliana by Laser Desorption / Ionization Mass Spectrometry Using

Silver Colloid as a Matrix Direct Profiling and Imaging of Epicuticular Waxes on Arabidopsis thaliana by Laser Desorption . Anal. Chem. 81, 2991–3000. https://doi.org/10.1021/ac802615r

Engewald, W., Teske, J., Efer, J., 1999. Programmed temperature vaporisers-based large volume injection in capillary gas chromatography. J. Chromatogr. A 842, 143–161. https://doi.org/10.1016/S0021-9673(99)00080-1

Fatland, B.L., Ke, J., Anderson, M.D., Mentzen, W.I., Cui, L.W., Allred, C.C., Johnston, J.L., Nikolau, B.J., Wurtele, E.S., 2002. Molecular Characterization of a Heteromeric ATP-Citrate Lyase That Generates Cytosolic Acetyl-Coenzyme A in Arabidopsis. Plant Physiol. 130, 740 LP-756. https://doi.org/10.1104/pp.008110

Fatland, B.L., Nikolau, B.J., Wurtele, E.S., 2005. Reverse Genetic Characterization of Cytosolic Acetyl-CoA Generation by ATP-Citrate Lyase in Arabidopsis. Plant Cell 17, 182 LP-203. https://doi.org/10.1105/tpc.104.026211

Feenstra, A.D., Alexander, L.E., Song, Z., Korte, A.R., Yandeau-Nelson, M.D., Nikolau, B.J., Lee, Y.J., 2017. Spatial mapping and profiling of metabolite distributions during germination. Plant Physiol. 174. https://doi.org/10.1104/pp.17.00652

Goodwin, S.M., Rashotte, A.M., Rahman, M., Feldmann, K.A., Jenks, M.A., 2005. Wax constituents on the inflorescence stems of double eceriferum mutants in Arabidopsis reveal complex gene interactions. Phytochemistry 66, 771–780. https://doi.org/10.1016/j.phytochem.2005.02.001

Hannoufa, A., McNevin, J., Lemieux, B., 1993. Epicuticular waxes of eceriferum mutants of Arabidopsis thaliana. Phytochemistry 33, 851–855. https://doi.org/https://doi.org/10.1016/0031-9422(93)85289-4

Haslam, T.M., Haslam, R., Thoraval, D., Pascal, S., Delude, C., Domergue, F., Fernández, A.M., Beaudoin, F., Napier, J.A., Kunst, L., Joubès, J., 2015. ECERIFERUM2-LIKE Proteins Have Unique Biochemical and Physiological Functions in Very-Long-Chain Fatty Acid Elongation. Plant Physiol. 167, 682–92. https://doi.org/10.1104/pp.114.253195

Jellings, A.J., Leech, R.M., 1982. The Importance of Quantitative Anatomy in the Interpretation of Whole Leaf Biochemistry in Species of Triticum, hordeum and Avena. New Phytol. 92, 39–48.

Jenks, M. a., Tuttle, H. a., Eigenbrode, S.D., Feldmann, K. a., 1995. Leaf Epicuticular Waxes of the Eceriferum Mutants in Arabidopsis. Plant Physiol. 108, 369–377. https://doi.org/10.1104/pp.108.1.369

Jun, J.H., Song, Z., Liu, Z., Nikolau, B.J., Yeung, E.S., Lee, Y.J., 2010. High-Spatial and High-Mass Resolution Imaging of Surface Metabolites of Arabidopsis thaliana by Laser Desorption-Ionization Mass Spectrometry Using Colloidal Silver. Anal. Chem. 82, 3255–3265. https://doi.org/10.1021/ac902990p

Jung, K.-H., Han, M.-J., Lee, D., Lee, Y.-S., Schreiber, L., Franke, R., Faust, A., Yephremov, A., Saedler, H., Kim, Y.-W., Hwang, I., An, G., 2006. Wax-deficient anther1; Is Involved in Cuticle and Wax Production in Rice Anther Walls and Is Required for Pollen Development. Plant Cell 18, 3015 LP-3032. https://doi.org/10.1105/tpc.106.042044

Kolattukudy, P.E., 1970. Biosynthesis of Cuticular Lipids. Annu. Rev. Plant Physiol. 21, 163–192. https://doi.org/10.1146/annurev.pp.21.060170.001115

Koornneef, M., Hanhart, C.J., Thiel, F., 1989. A Genetic and Phenotypic Description of Eceriferum (cer) Mutants in Arabidopsis thaliana. J. Hered. 80, 118–122.

Kunst, L., Samuels, L., 2009. Plant cuticles shine: advances in wax biosynthesis and export. Curr. Opin. Plant Biol. 12, 721–727. https://doi.org/10.1016/j.pbi.2009.09.009

Lee, S.B., Suh, M.C., 2013. Recent advances in cuticular wax biosynthesis and its regulation in arabidopsis. Mol. Plant 6, 246–249. https://doi.org/10.1093/mp/sss159

Lee, Y.J., Perdian, D.C., Song, Z., Yeung, E.S., Nikolau, B.J., Jin, L.Y., C., P.D., Zhihong, S., S., Y.E., J., N.B., 2012. Use of mass spectrometry for imaging metabolites in plants. Plant J. 70, 81–95. https://doi.org/10.1111/j.1365-313X.2012.04899.x

Leide, J., Hildebrandt, U., Reussing, K., Riederer, M., Vogg, G., 2007. The Developmental Pattern of Tomato Fruit Wax Accumulation and Its Impact on Cuticular Transpiration Barrier Properties: Effects of a Deficiency in a β-Ketoacyl-Coenzyme A Synthase (LeCER6). Plant Physiol. 144, 1667 LP-1679. https://doi.org/10.1104/pp.107.099481

Millar, A.A., Clemens, S., Zachgo, S., Giblin, E.M., Taylor, D.C., Kunst, L., 1999. CUT1, an Arabidopsis Gene Required for Cuticular Wax Biosynthesis and Pollen Fertility, Encodes a Very-Long-Chain Fatty Acid Condensing Enzyme. Plant Cell 11, 825 LP-838. https://doi.org/10.1105/tpc.11.5.825

Pascal, S., Bernard, A., Sorel, M., Pervent, M., Vile, D., Haslam, R.P., Napier, J.A., Lessire, R., Domergue, F., Joubès, J., 2013. The Arabidopsis cer26 mutant, like the cer2 mutant, is specifically affected in the very long chain fatty acid elongation process. Plant J. 73, 733–746. https://doi.org/10.1111/tpj.12060

Post-Beittenmiller, D., 1996. Biochemistry and Molecular Biology of Wax Production in Plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 405–430. https://doi.org/10.1146/annurev.arplant.47.1.405

Samuels, L., Kunst, L., Jetter, R., 2008. Sealing plant surfaces: cuticular wax formation by epidermal cells. Annu. Rev. Plant Biol. 59, 683–707. https://doi.org/10.1146/annurev.arplant.59.103006.093219

Schnable PS, Stinard PS, Wen TJ, Heinen S, Weber D, Zhang L, Hansen JD, and Nikolau BJ (1994) The genetics of cuticular wax biosynthesis, Maydica 39, 279-287

Smyth, D.R., 1990. Early Flower Development in Arabidopsis. Plant Cell 2, 755–767. https://doi.org/10.1105/tpc.2.8.755

Vogg, G., Fischer, S., Leide, J., Emmanuel, E., Jetter, R., Levy, A.A., Riederer, M., 2004. Tomato fruit cuticular waxes and their effects on transpiration barrier properties: functional characterization of a mutant deficient in a very- long- chain fatty acid β- ketoacyl- CoA synthase. J. Exp. Bot. 55, 1401–1410.

Wakao, S., Andre, C., Benning, C., 2008. Functional Analyses of Cytosolic Glucose-6-Phosphate Dehydrogenases and Their Contribution to Seed Oil Accumulation in Arabidopsis. Plant Physiol. 146, 277 LP-288. https://doi.org/10.1104/pp.107.108423

Wakao, S., Benning, C., 2005. Genome-wide analysis of glucose-6-phosphate dehydrogenases in Arabidopsis. Plant J. 41, 243–256. https://doi.org/10.1111/j.1365-313X.2004.02293.x

Wheeler, M.C.G., Tronconi, M.A., Drincovich, M.F., Andreo, C.S., Flügge, U.-I., Maurino, V.G., 2005. A Comprehensive Analysis of the NADP-Malic Enzyme Gene Family of Arabidopsis. Plant Physiol. 139, 39 LP-51. https://doi.org/10.1104/pp.105.065953

Wang, X., Guan, Y., Zhang, D., Dong, X., Tian, L., Qu, L.Q., 2017. A β-Ketoacyl-CoA Synthase Is Involved in Rice Leaf Cuticular Wax Synthesis and Requires a CER2-LIKE Protein as a Cofactor. Plant Physiol. 173, 944 LP-955. https://doi.org/10.1104/pp.16.01527

Wilson, B., Wylie, P.L., Klee, M.S., 2000. Large Volume Injection for Gas Chromatography Using a PTV Inlet I, 1–6.

Xia, Y., Nikolau, B.J., Schnable, P.S., 1996. Cloning and characterization of CER2, an Arabidopsis gene that affects cuticular wax accumulation. Plant Cell 8, 1291–1304. https://doi.org/10.1105/tpc.8.8.1291

Yin, Y., Ashihara, H., 2008. Expression of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase isoform genes in suspension-cultured Arabidopsis thaliana cells. Z. Naturforsch. C. 63, 713. https://doi.org/10.1515/znc-2008-9-1017

Yu, D., Ranathunge, K., Huang, H., Pei, Z., Franke, R., Schreiber, L., He, C., 2008. Wax Crystal-Sparse Leaf1 encodes a β–ketoacyl CoA synthase involved in biosynthesis of cuticular waxes on rice leaf. Planta 228, 675–685. https://doi.org/10.1007/s00425-008-0770-9

Figures

Figure 1. Arabidopsis flower development. Stage A are flowers with closed buds; Stage B flowers are identified by the first emergence of petals from the bud; Stage C, the emerging petals are perpendicular to the flower axis, and at Stage D the flower recloses and the developing silique begins to emerge.

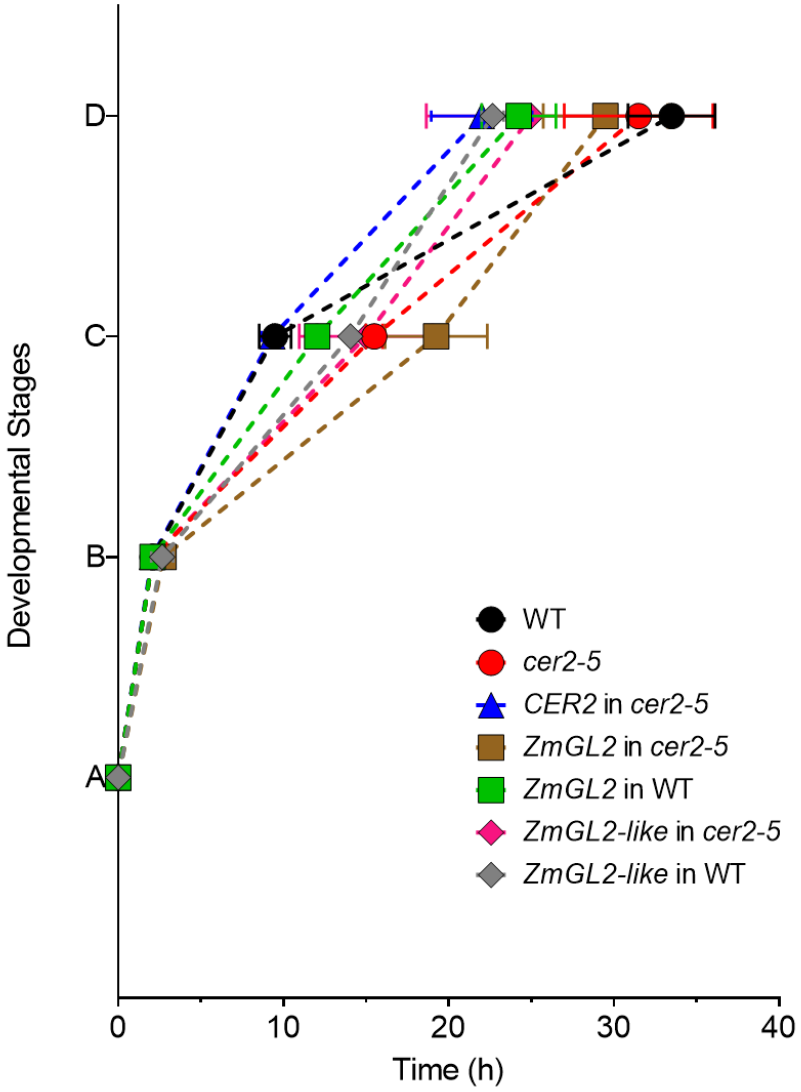


Figure 2. Timing of Arabidopsis flower development. Timing of transitions from Stage A to Stages, B, C, and D of non-transgenic and transgenic wild-type and *cer2-5* mutant flowers expressing either the *CER2* gene or the maize *Gl2* or *Gl2-like* genes. The data is the average of 2-4 replicates ± standard error, determined from time-lapse videos.

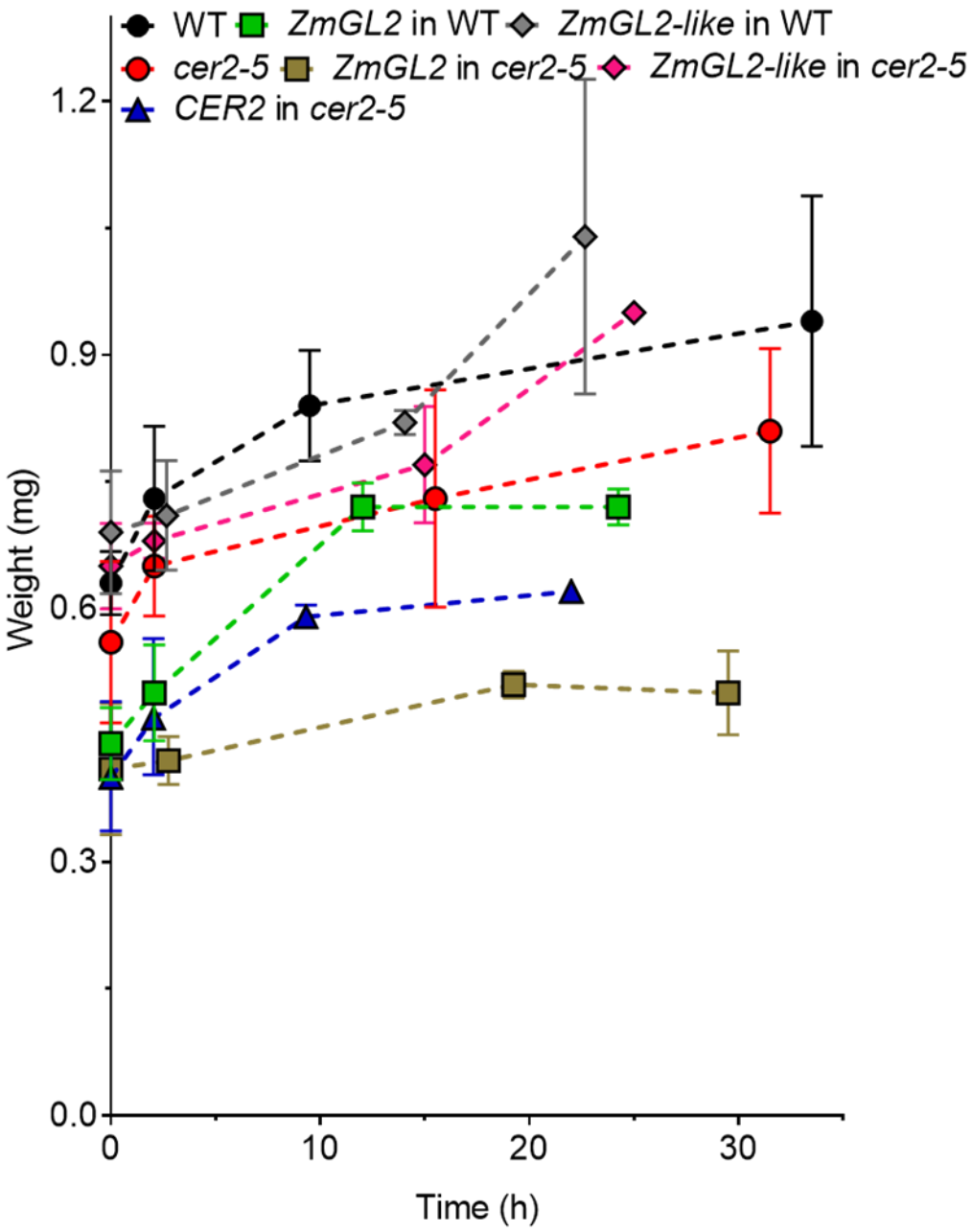


Figure 3. Change in biomass with flower development. Fresh weight of flowers ± standard deviation is plotted against time taken for Arabidopsis flowers over-expressing maize transgenes *Gl2* or *Gl2-like* to develop. Biomass of Arabidopsis flowers. The fresh weight of 6-13 pooled flowers of the indicated genotypes were weighed at each of the 4 flower stages defined in Figures 1 and 2. The data is the average of 2-3 determinations ± standard deviation.

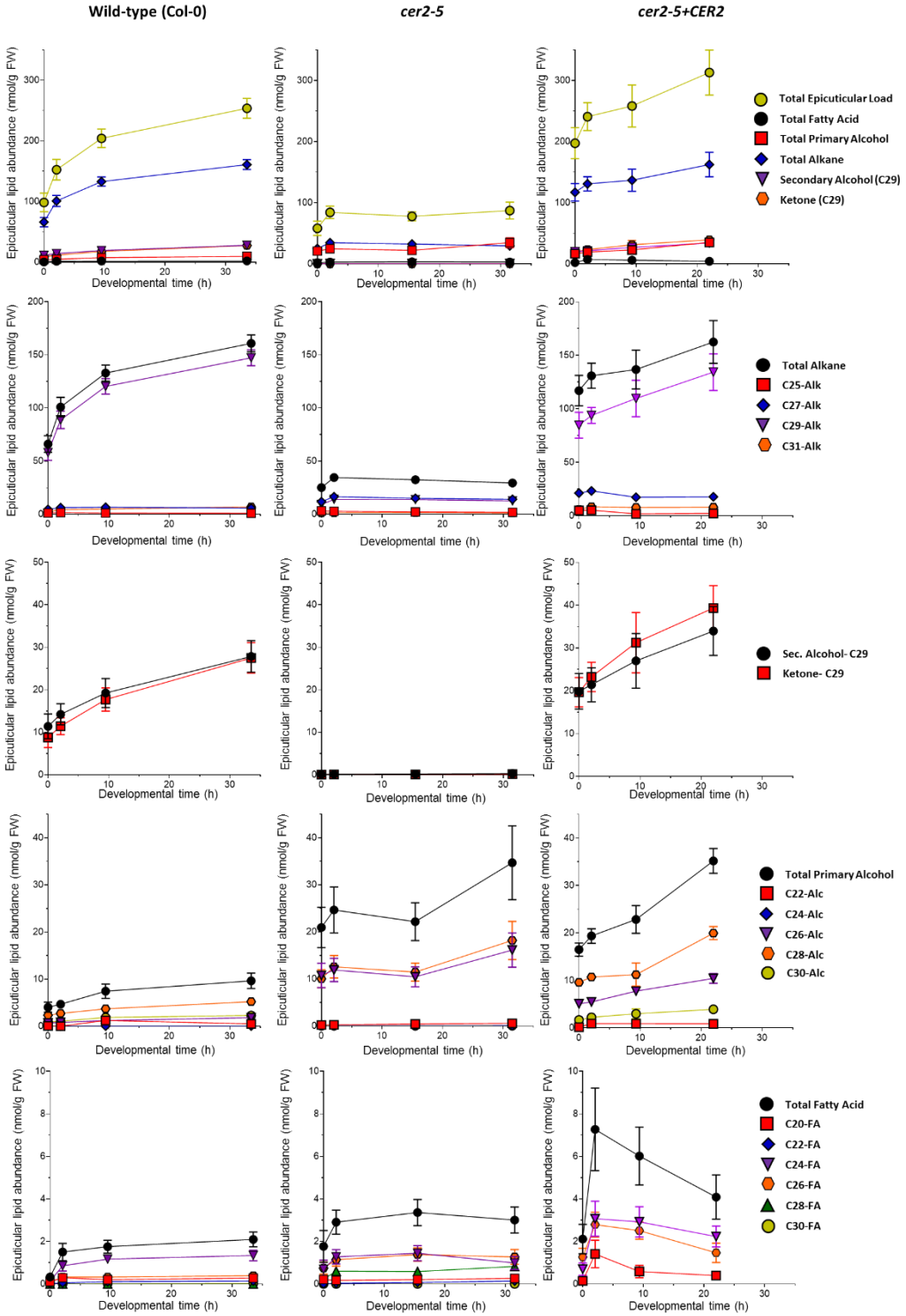


Figure 4. Epicuticular surface lipid accumulation on developing Arabidopsis flowers. Epicuticular surface lipids were extracted from individual Arabidopsis flowers of wild-type, *cer2-5* mutant, and *CER2* overexpressed in the mutant, and analyzed using a GC equipped with a LVI-PTV inlet injector (see Materials and Methods). The data represents the average ± standard error, of two independent experiments for the wild-type and *cer2-5* mutant arranged in order of most abundant metabolites.

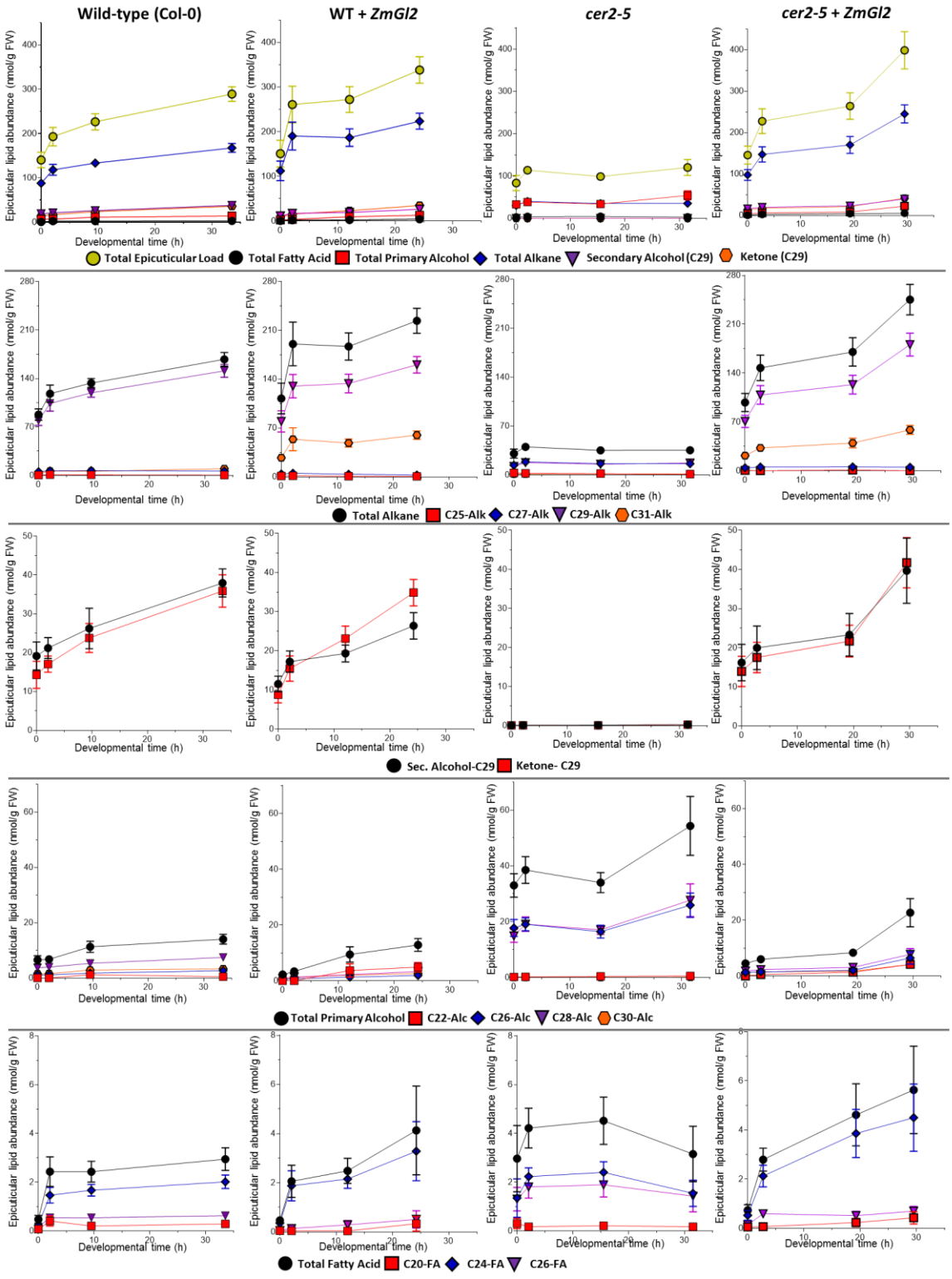


Figure 5. Epicuticular surface lipid trends on developing Arabidopsis flowers. Epicuticular surface lipids were extracted from individual flowers of genotypes: wild-type (Col-0), *cer2-5* mutant, and either of these lines expressing maize *Gl2* transgene. The data represent average ± standard error of 6 replicates arranged in order of highest abundance.

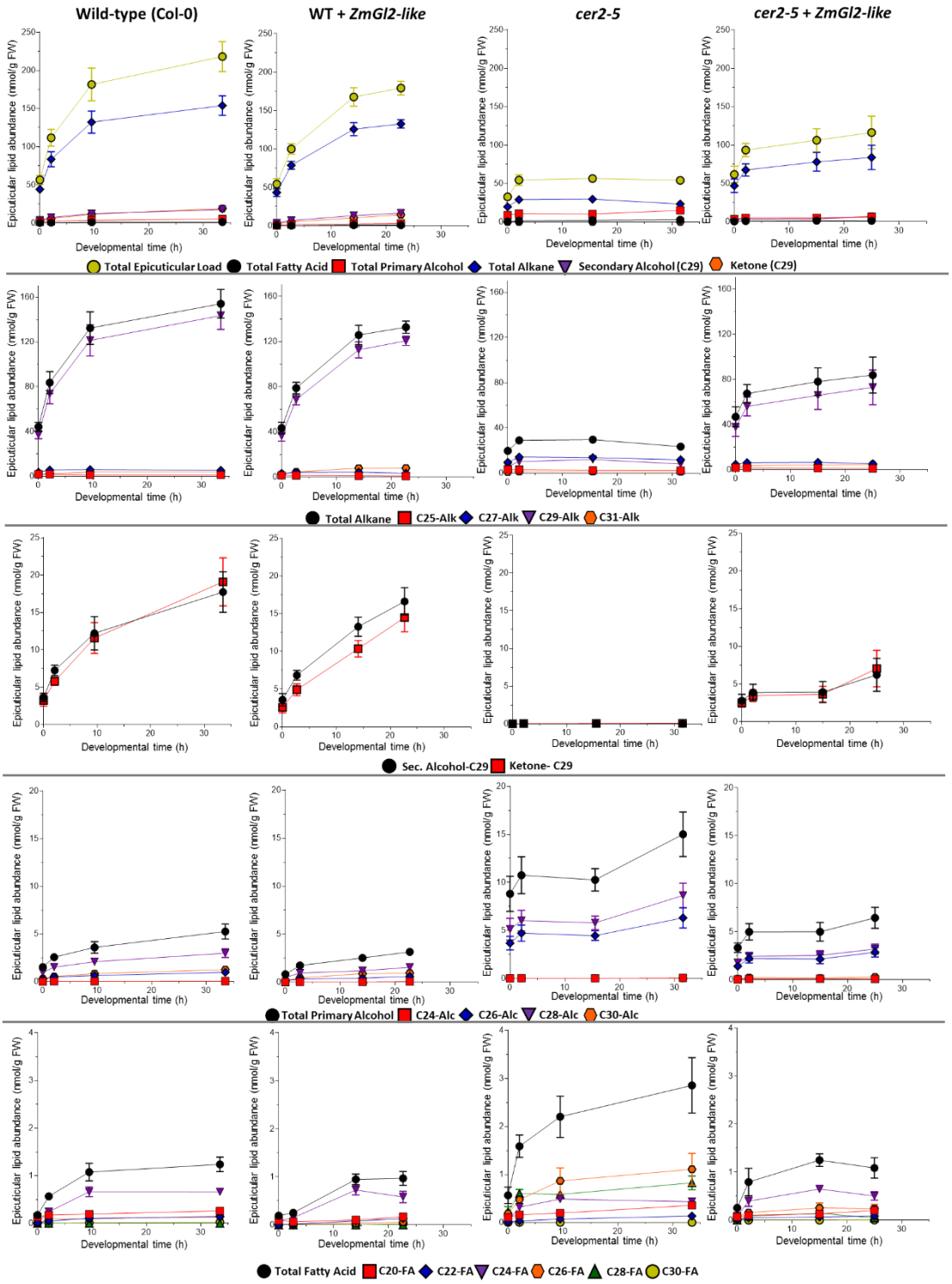


Figure 6. Epicuticular surface lipid trends on developing Arabidopsis flowers. Epicuticular surface lipids were extracted from individual flowers of genotypes: wild-type (Col-0), *cer2-5* mutant, and either of these lines expressing maize *Gl2-like* transgene. The data represent average ± standard error of 6 replicates arranged in order of highest abundance.

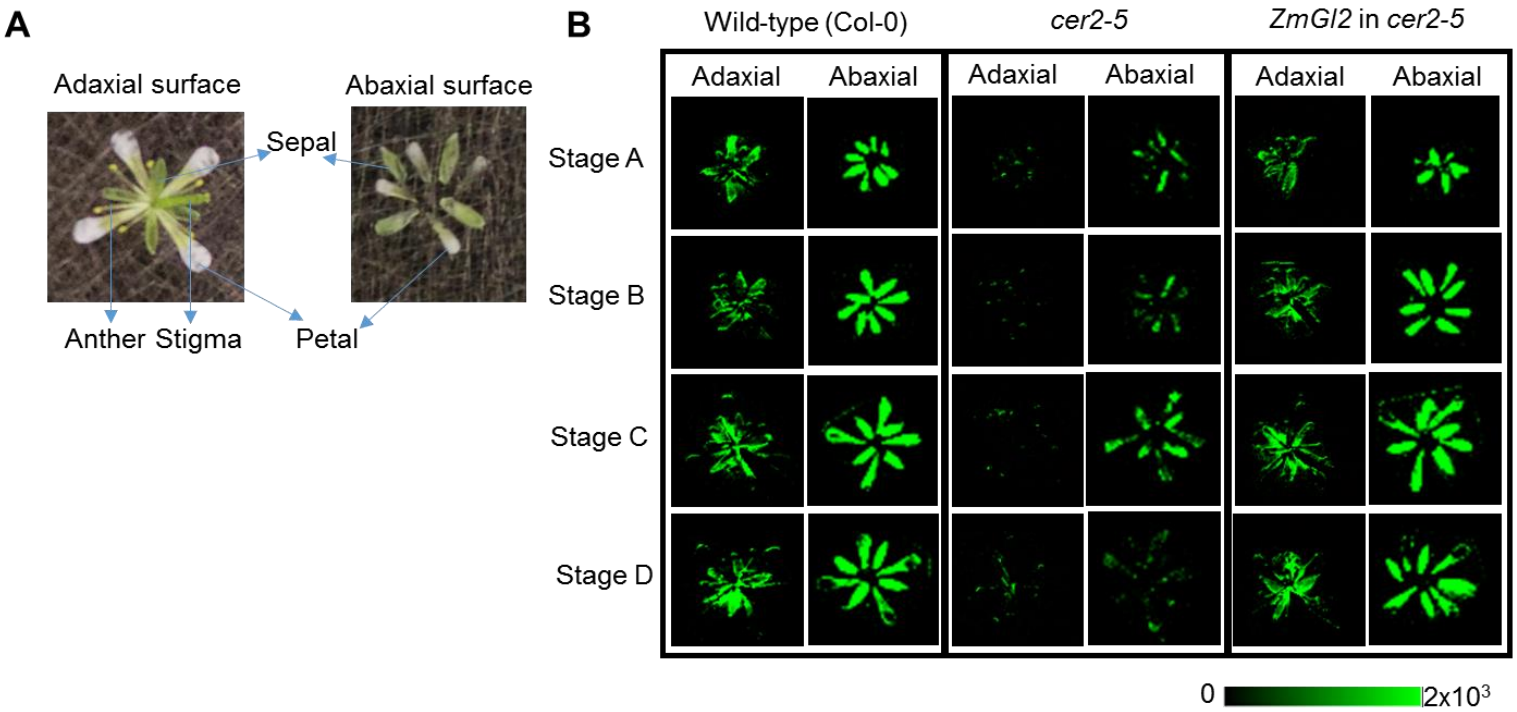


Figure 7. Spatial distribution of nonacosane on Arabidopsis flowers. A) Visual images of the adaxial and abaxial surfaces of Arabidopsis flowers positioned on a MALDI plate ready for mass spectrometric imaging. B) The distribution of nonacosane (C29 alkane) on the adaxial and abaxial floral surfaces at four stages of development (Stages A to D) of the indicated genotypes. MSI absolute intensity scale bar is color coded: green is maximum signal and black is minimum signal.

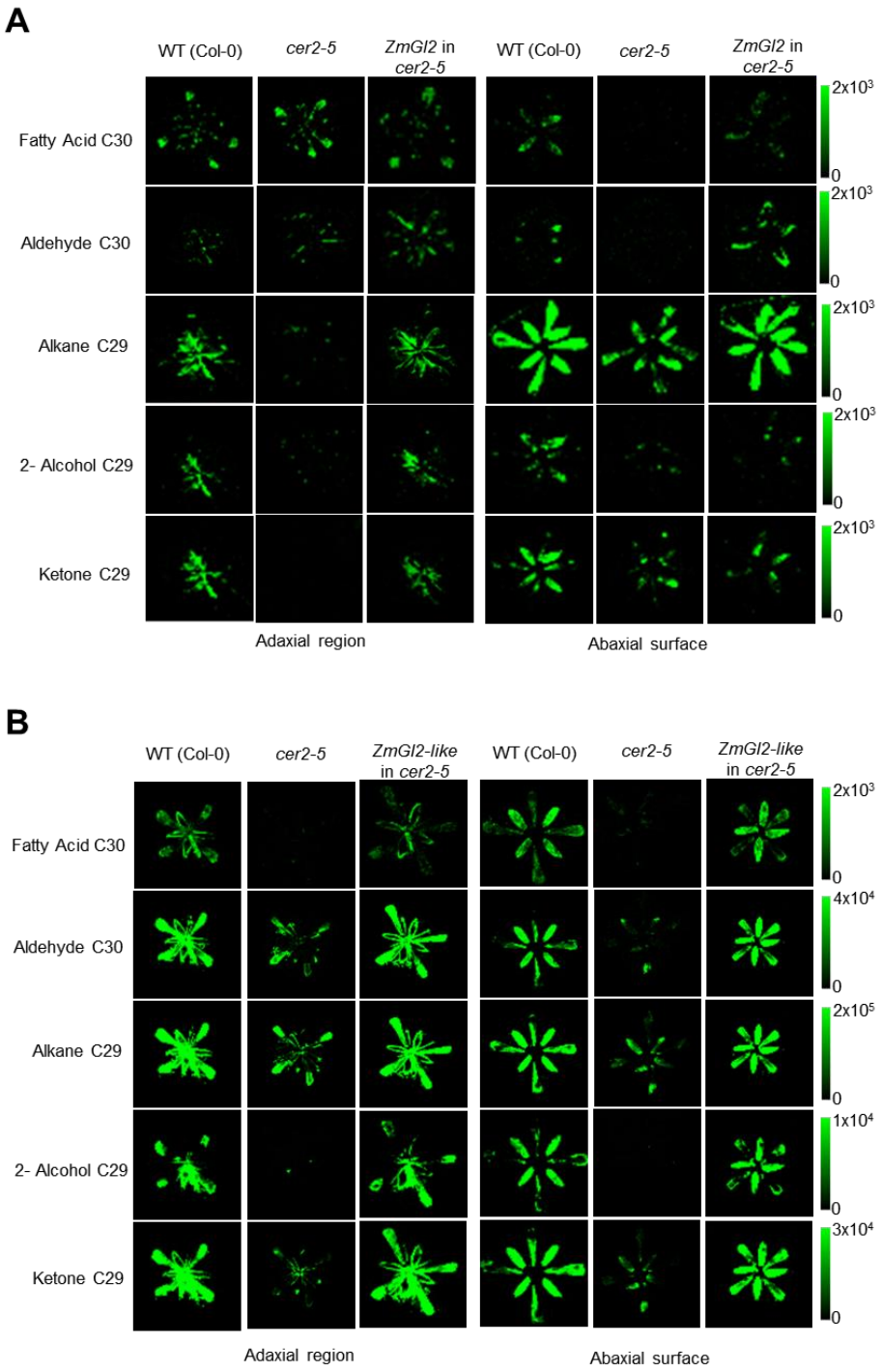


Figure 8. Spatial distribution of epicuticular lipids on the adaxial and abaxial surfaces of Arabidopsis flowers. Two independent experiments that visualize the effect of the transgenic expression of *Gl2* (A) or *Gl2-like* (B) on the spatial distribution of C30-fatty acid and alkyl-derivatives on the two surfaces of Arabidopsis flowers. The genotype of each flower is indicated and MSI absolute intensity scale bar is color coded: green is maximum signal and black is minimum signal.

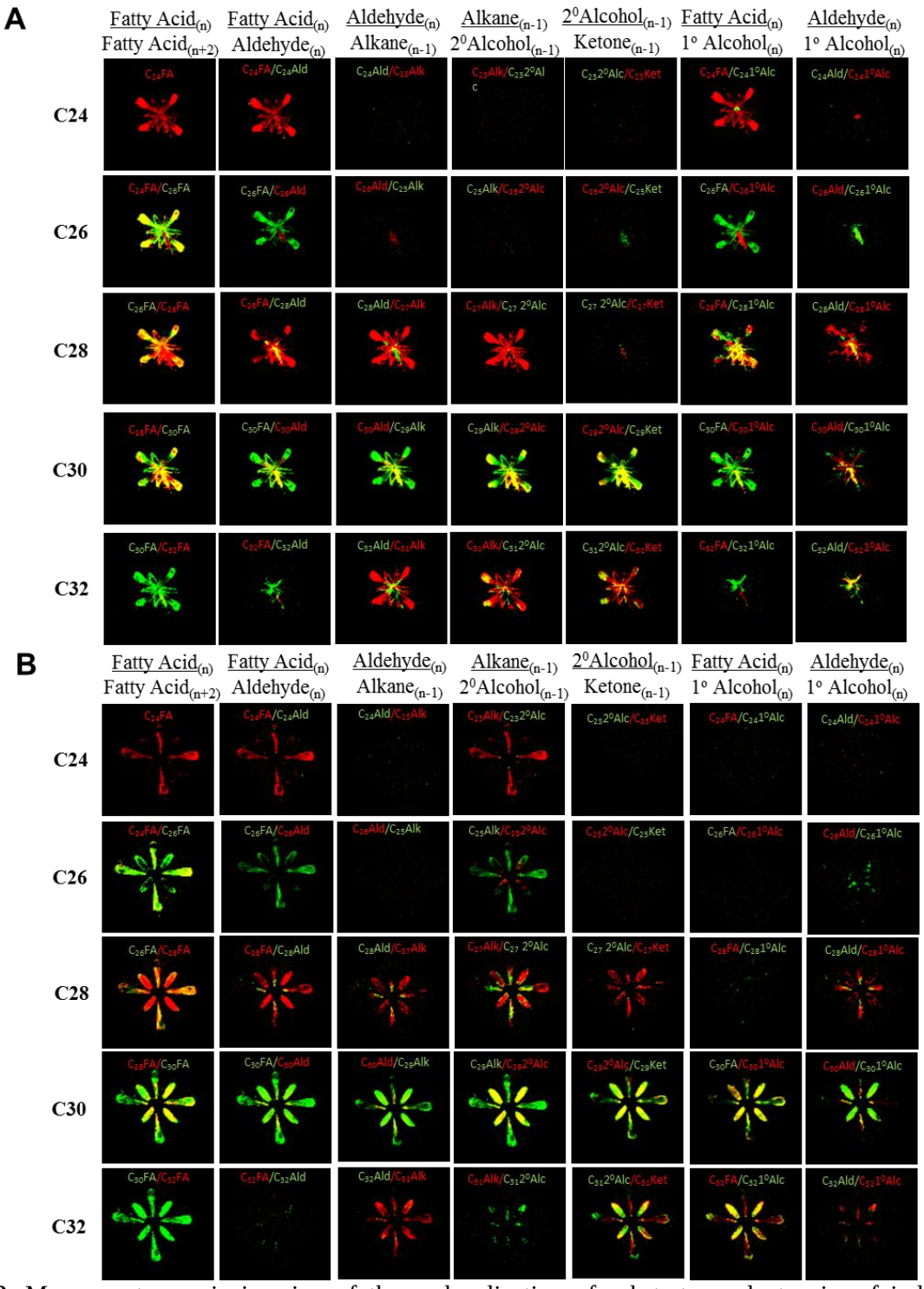
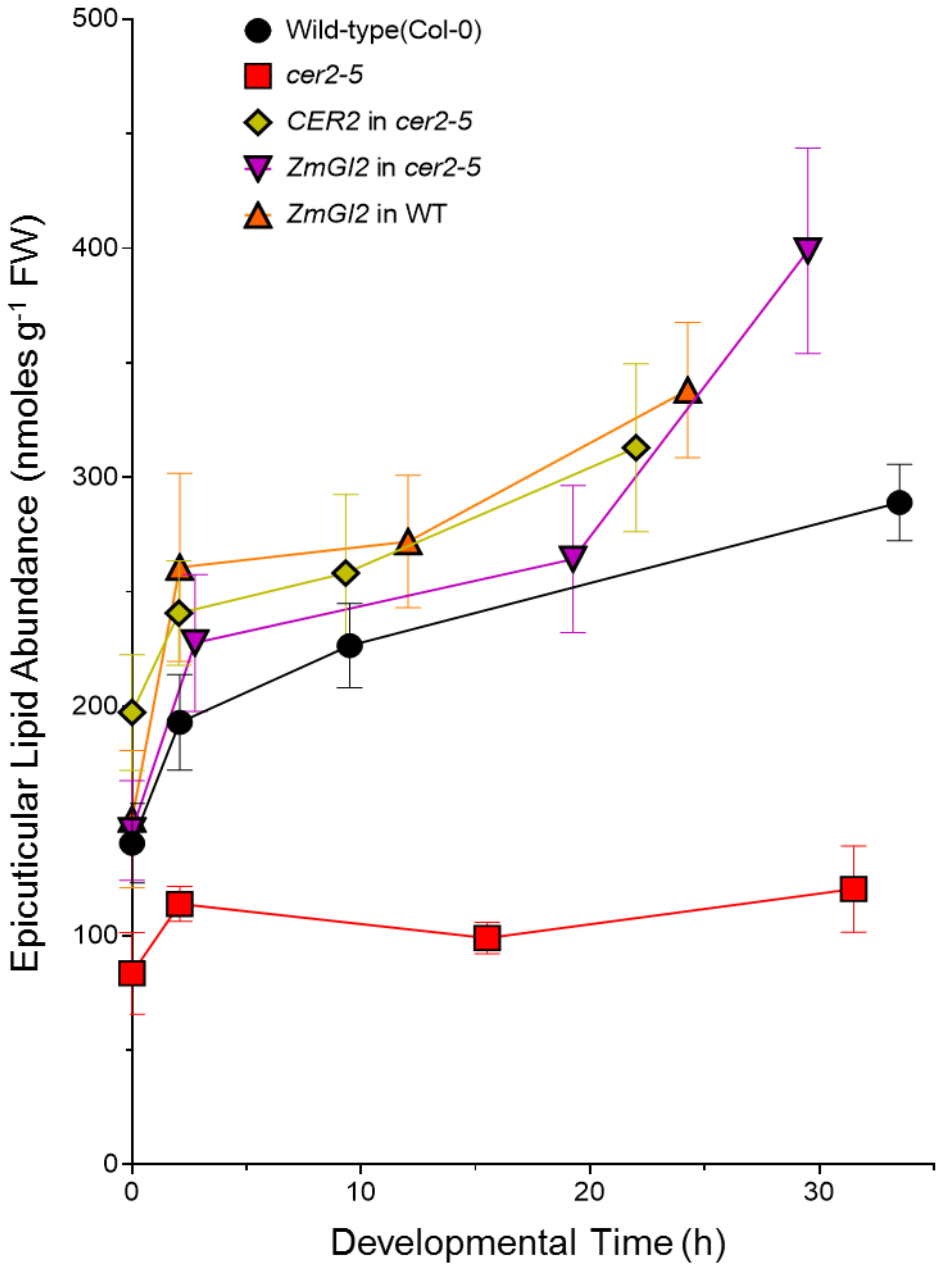
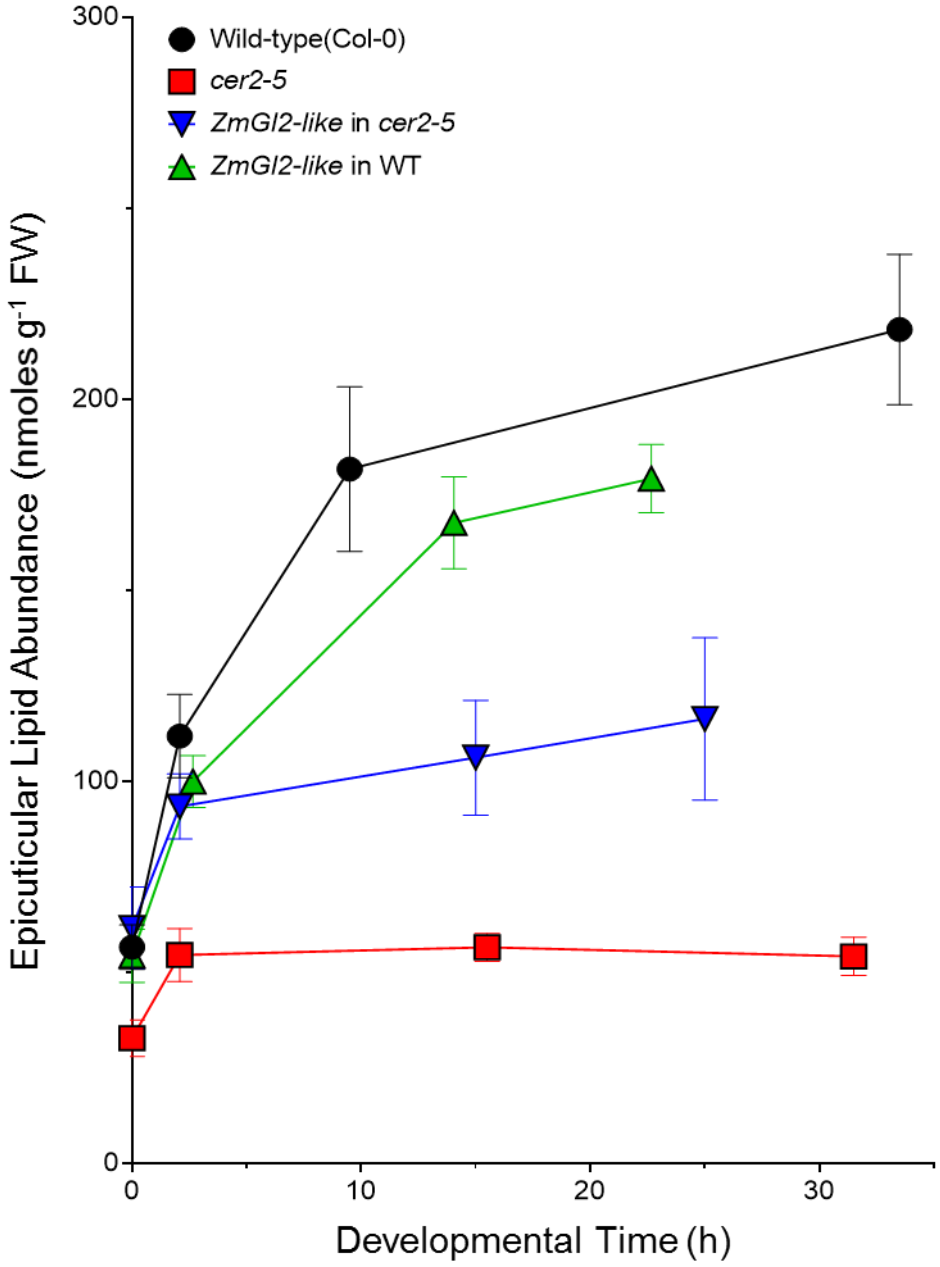


Figure 9. Mass spectroscopic imaging of the co-localization of substrate-product pairs of individual reactions of the epicuticular lipid biosynthesis pathway. Each image is a fusion of two false color-coded MSI images, green for the substrate, and red for the product of a single metabolic reaction that contributes to epicuticular lipid biosynthesis pathway. In these fused images, yellow represents the spatial zone where these substrate-product pairs co-localize, indicating the location of the underlying tissue/cells where this metabolic interconversion could be occurring. The distribution of C24 to C33-fatty acid and their alkyl derivatives on the adaxial (A) and abaxial (B) floral surfaces is indicated for wild-type Col-0.

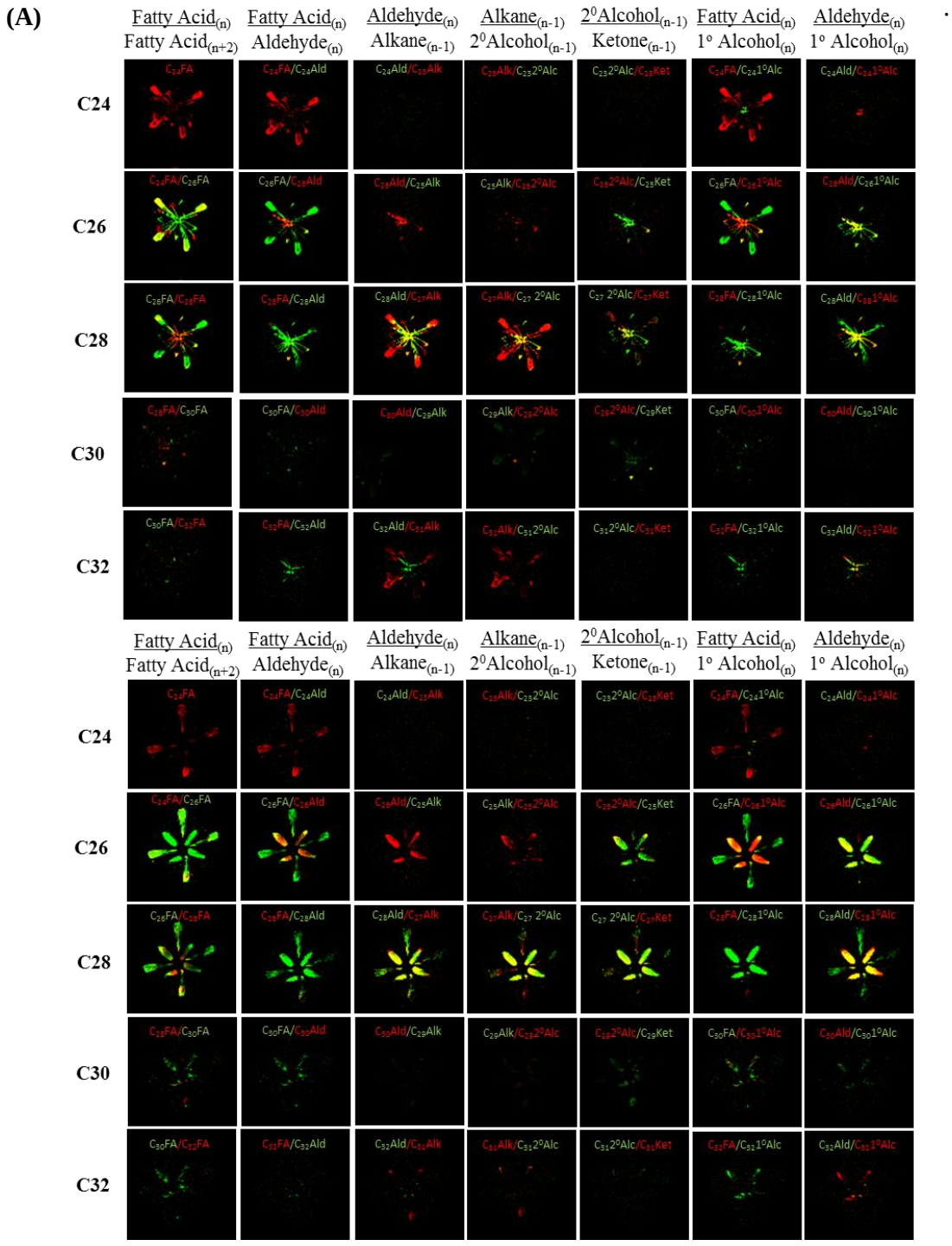
Supplemental Figures and Tables



Supplemental Figure 1. Epicuticular surface lipid accumulation on developing Arabidopsis flowers. Epicuticular surface lipids were extracted from individual Arabidopsis flowers of the indicated genotypes, wild-type, *cer2-5*, and either of the two backgrounds expressing maize *Gl2* transgene, and analyzed using a GC equipped with a LVI-PTV inlet injector (see Materials and Methods). The data represent average $\pm$ standard error of 6 replicates.

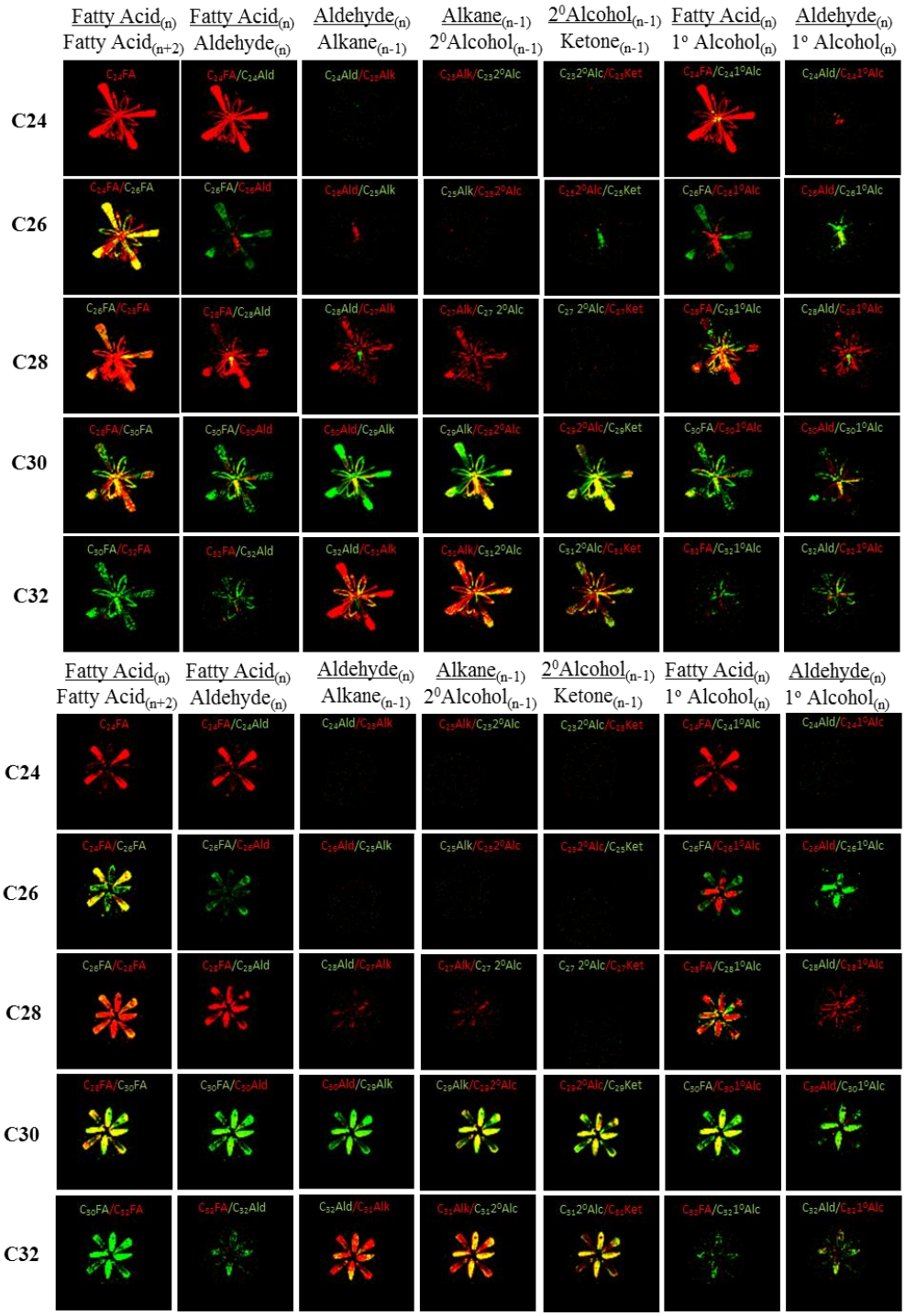


Supplemental Figure 2 Epicuticular surface lipid accumulation on developing Arabidopsis flowers. Epicuticular surface lipids were extracted from individual Arabidopsis flowers of the indicated genotypes, wild-type, *cer2-5*, and either of the two backgrounds expressing maize *Gl2-like* transgene and analyzed using a GC equipped with a LVI-PTV inlet injector (see Materials and Methods). The data represent average ± standard error of 6 replicates.



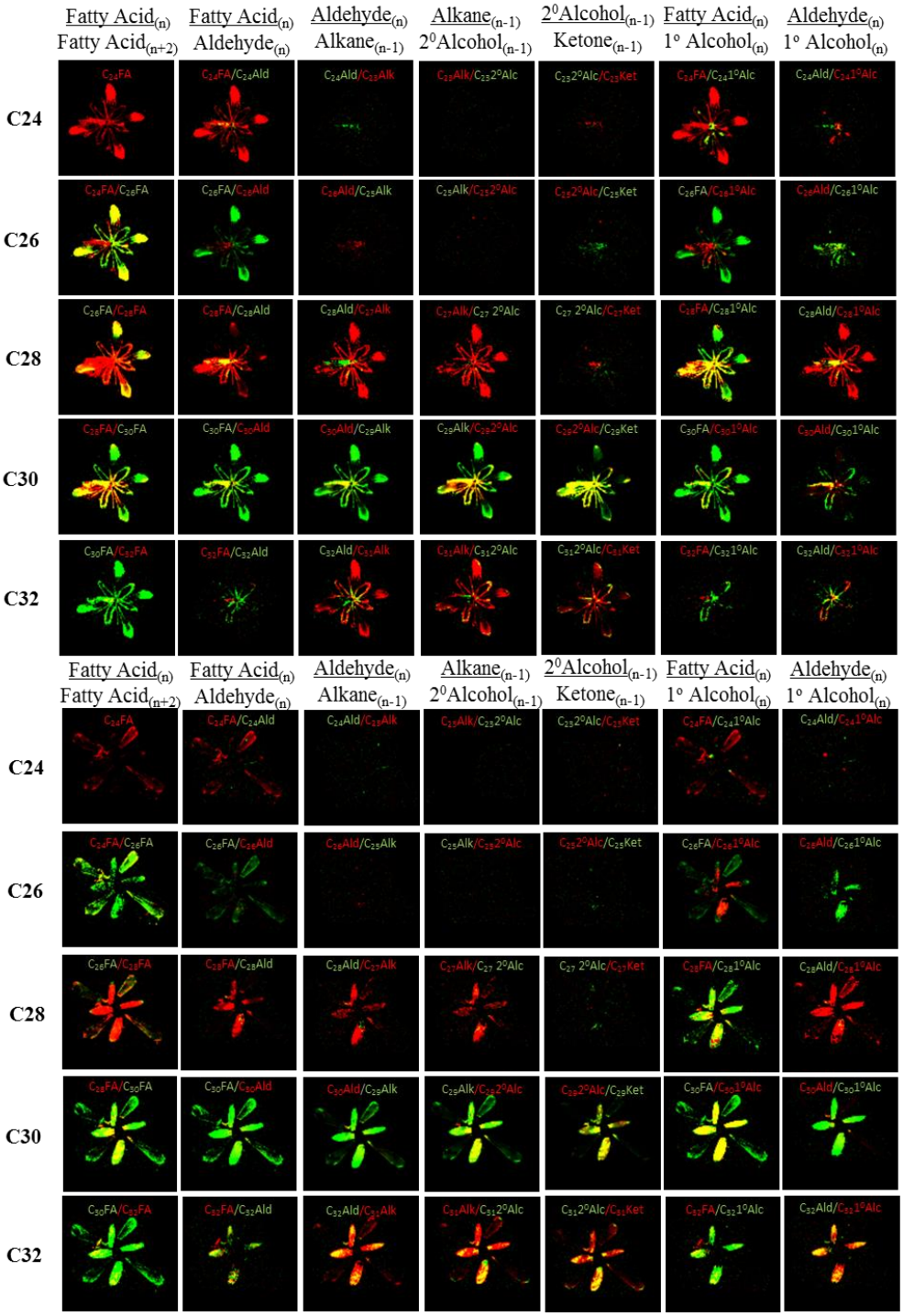
Supplemental Figure 3. Mass spectroscopic imaging of the co-localization of substrate-product pairs of individual reactions of the epicuticular lipid biosynthesis pathway. Each image is a fusion of two false color-coded MSI images, green for the substrate, and red for the product of a single metabolic reaction that contributes to epicuticular lipid biosynthesis pathway as denoted for C24- to -C32 fatty acids and their alkyl derivatives. In these fused images, yellow represents the spatial zone where these substrate-product pairs co-localize, indicating the location of the underlying tissue/cells where this metabolic interconversion could

(B)

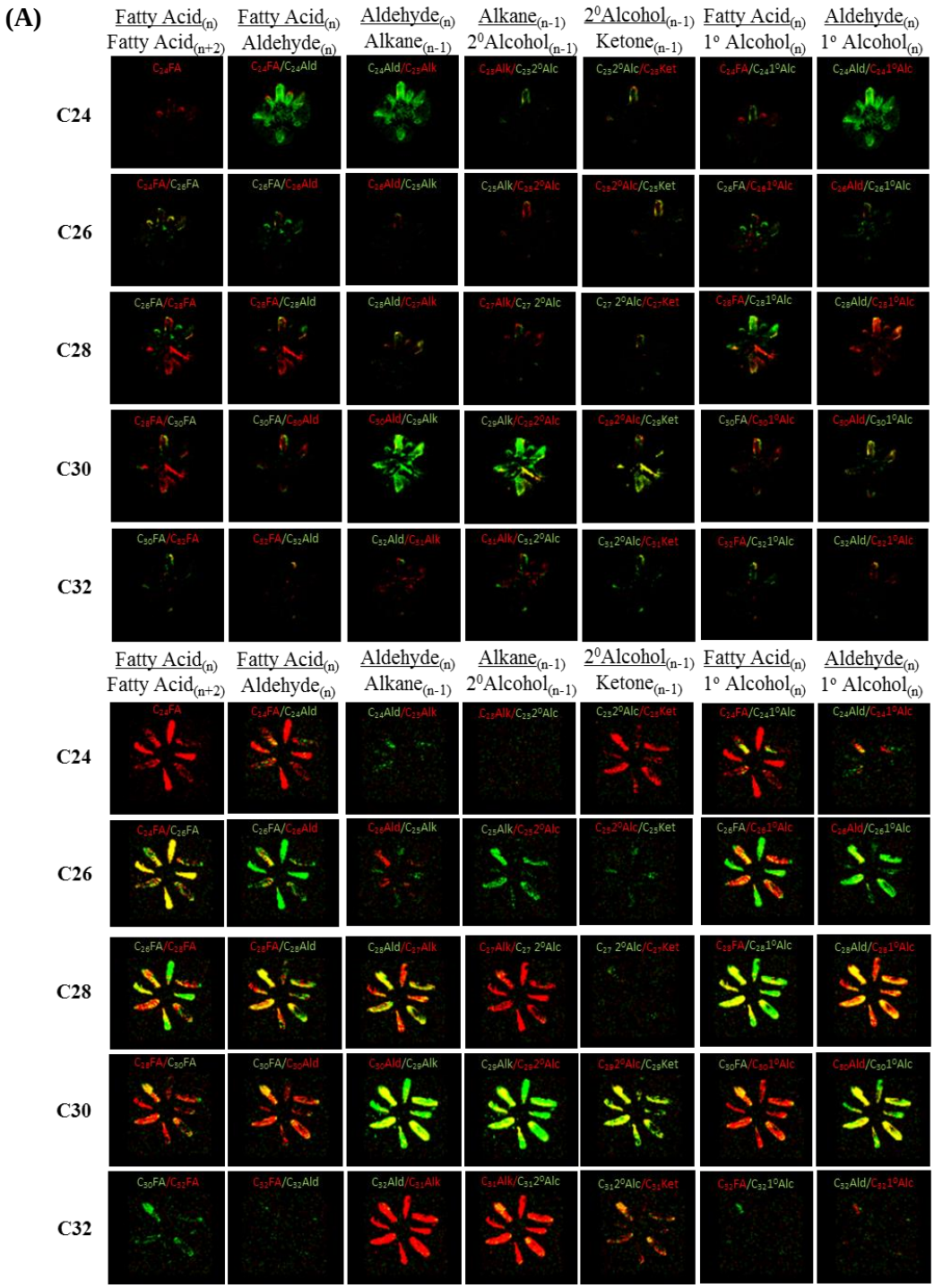


be occurring. Images are indicated for the adaxial (top image) and abaxial (lower image) surfaces of Arabidopsis flowers at Stage C development for the following genotypes: (A) *cer2-5* mutant, (B) *Gl2-like* transgene expressed in *cer2-5* mutant background, (C) *Gl2-like* transgene expressed in wild-type background

(C)

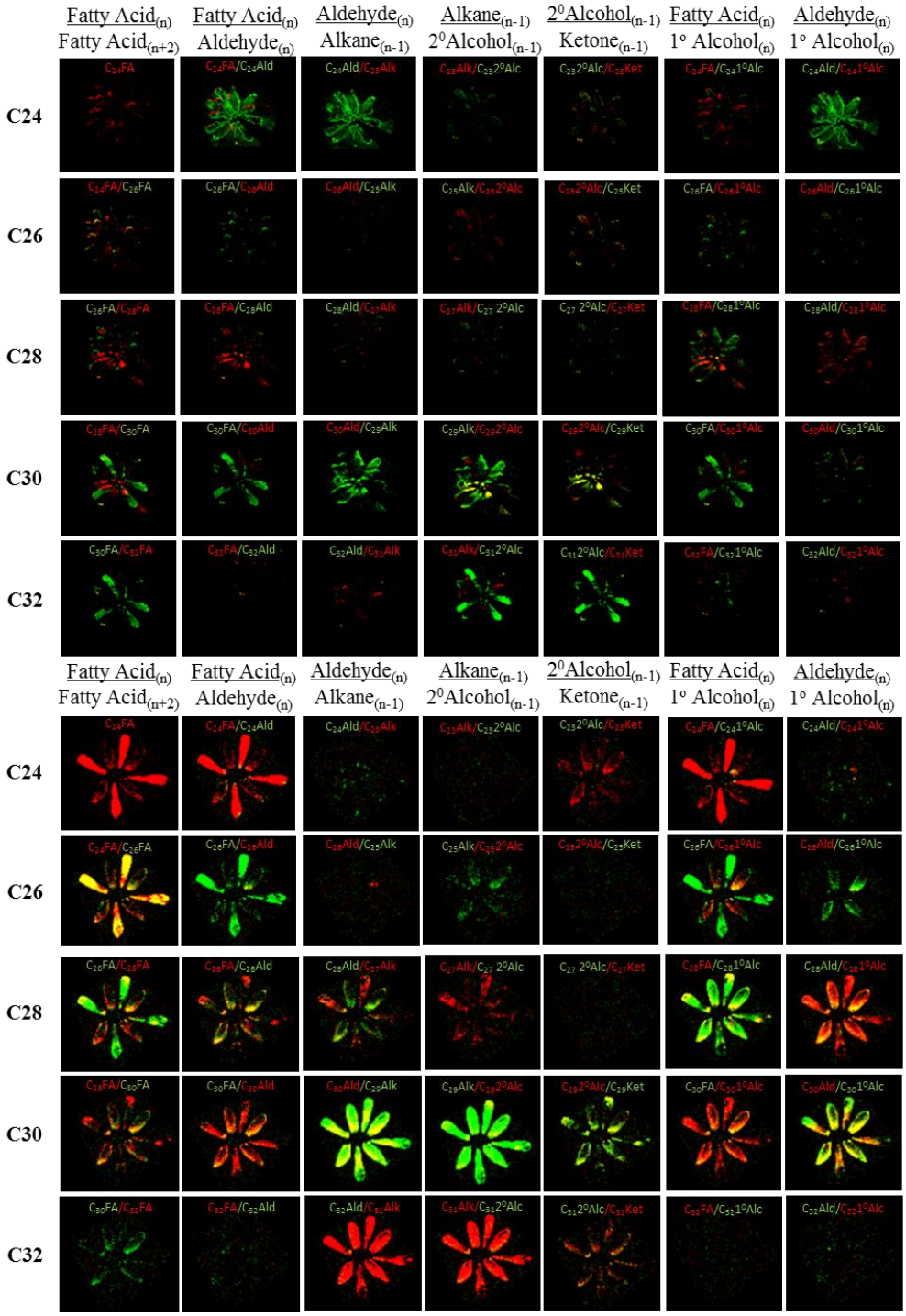


Supplemental Figure 3 cont'd

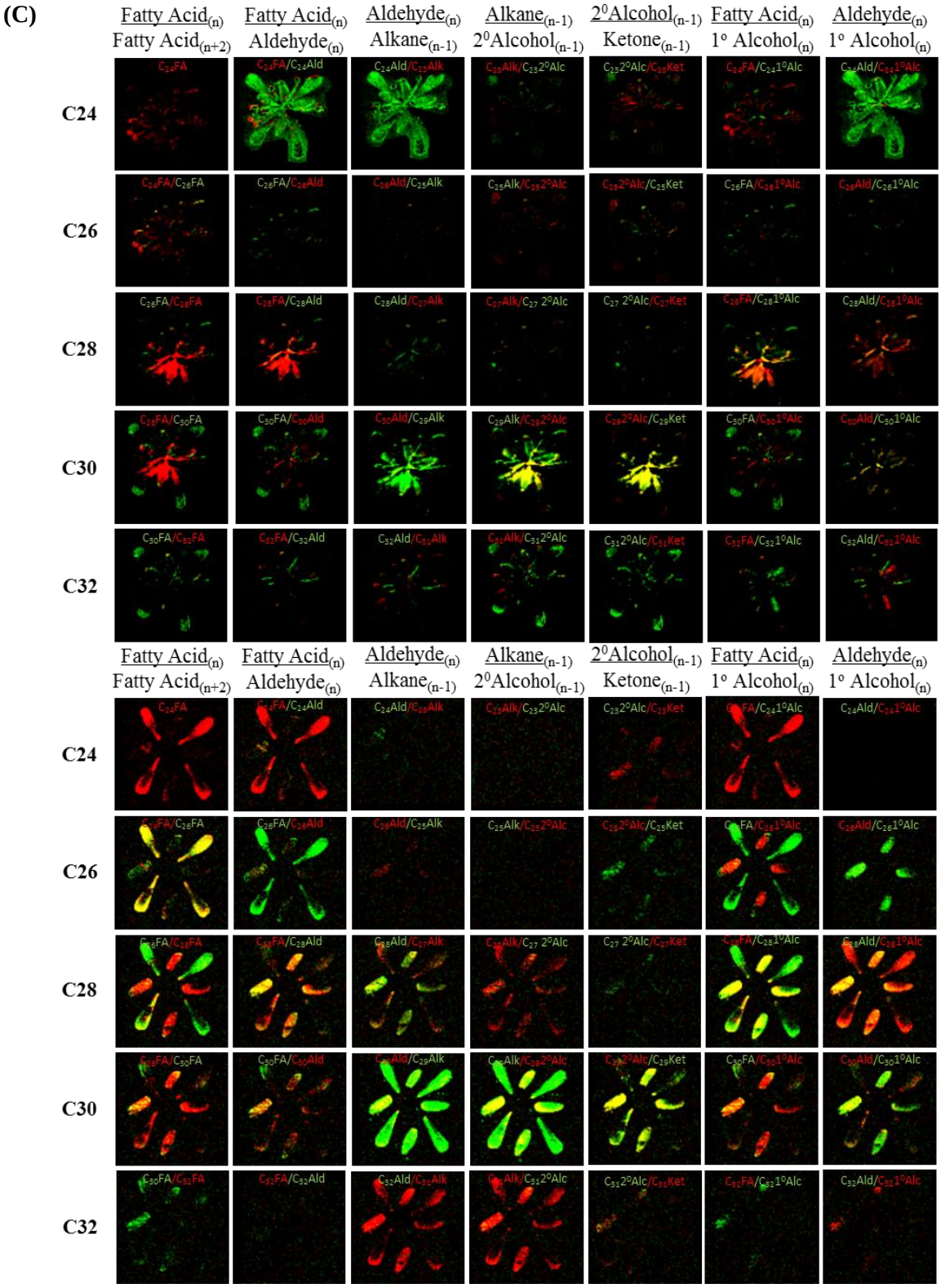


Supplemental Figure 4. Mass spectroscopic imaging of the co-localization of substrate-product pairs of individual reactions of the epicuticular lipid biosynthesis pathway. Each image is a fusion of two false color-coded MSI images, green for the substrate, and red for the product of a single metabolic reaction that contributes to epicuticular lipid biosynthesis pathway occurring during four stages of flower development (Stage A to D). In these fused images, yellow represents the spatial zone where these substrate-product pairs co-localize, indicating the location of the underlying tissue/cells where this metabolic interconversion could be occurring. Images are indicated for the adaxial (top image) and abaxial (lower image) surfaces of

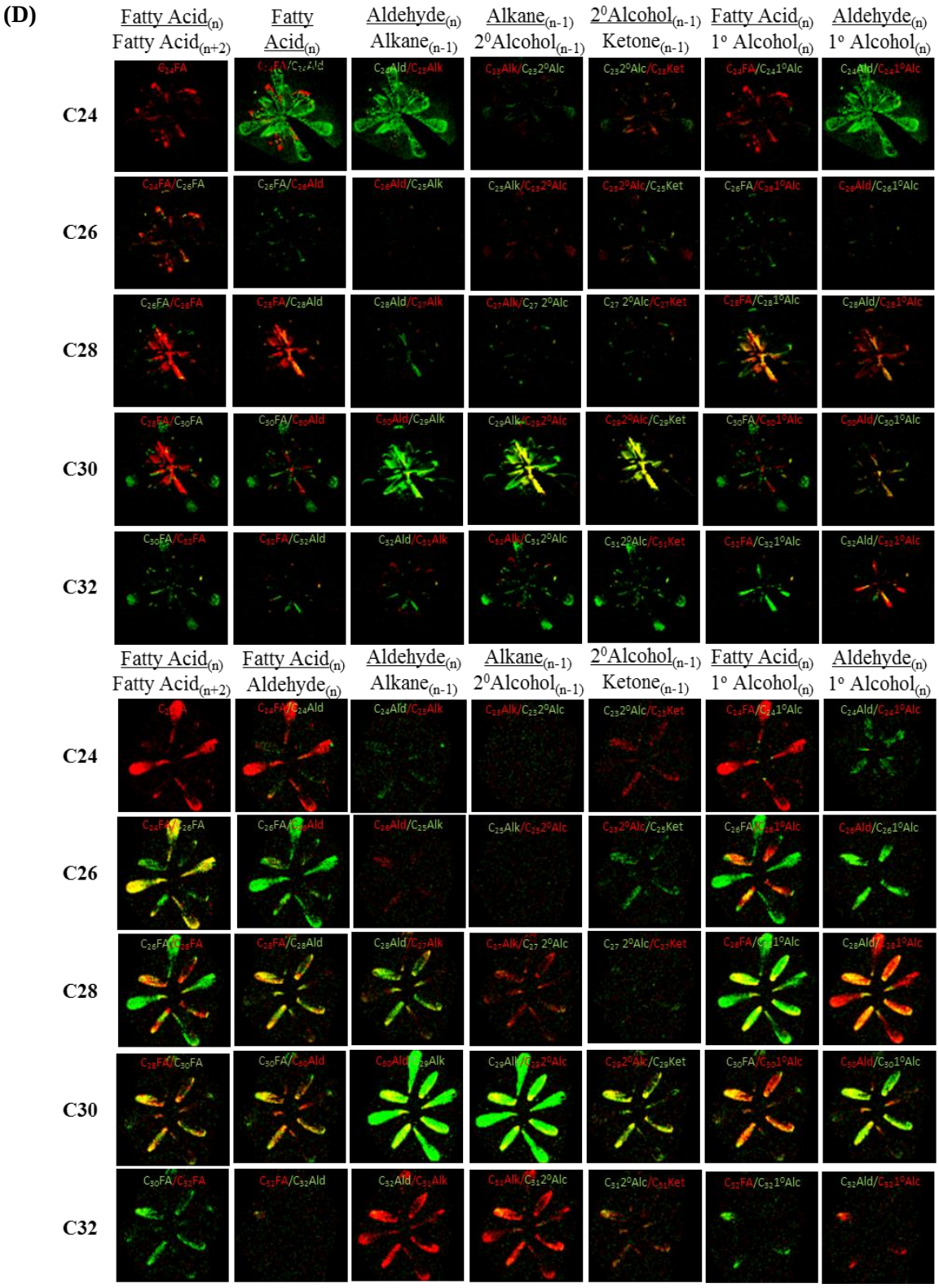
(B)



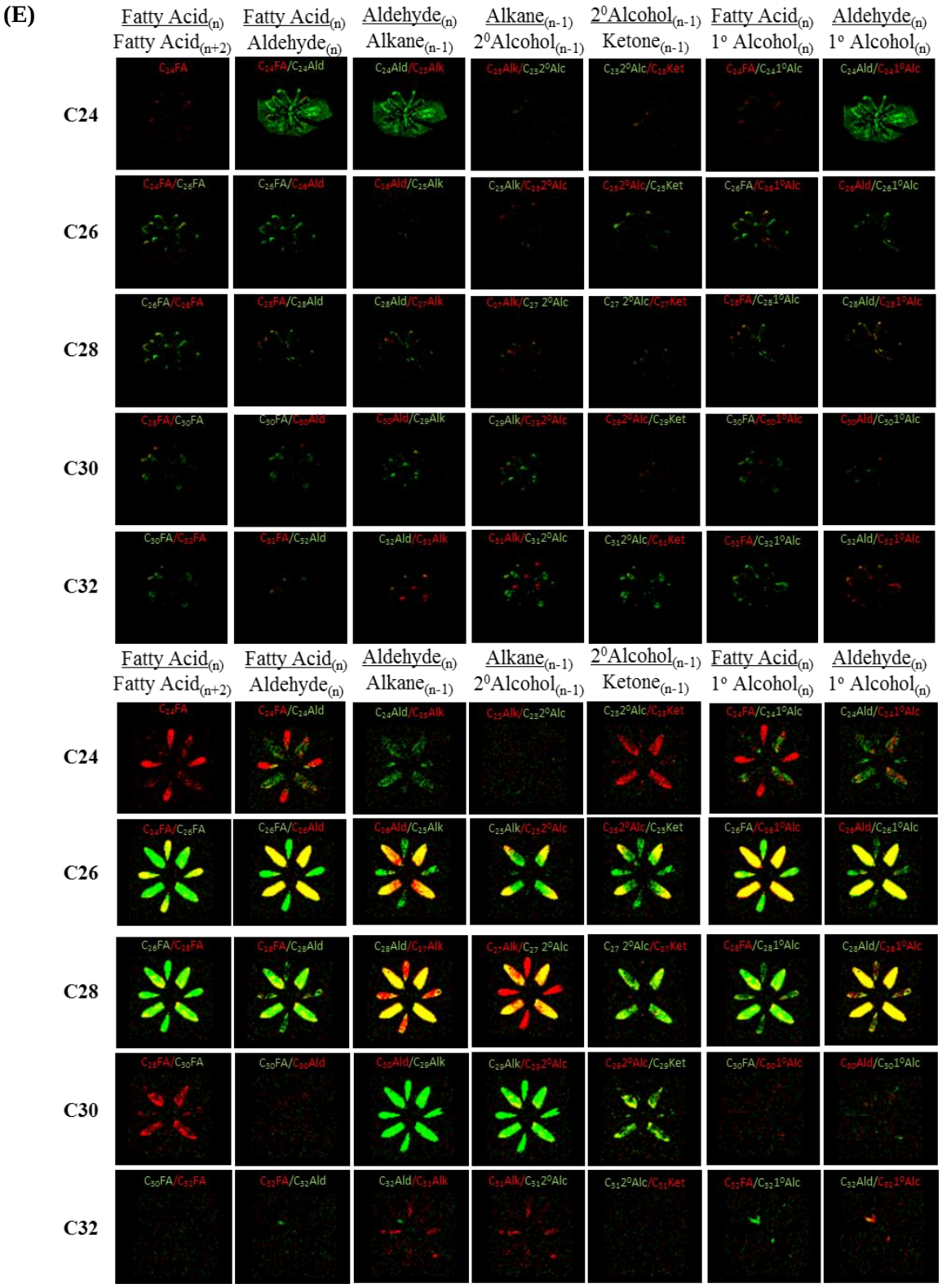
Arabidopsis flowers of genotypes: (A-D) Wild-type (Col-0) Stages A-D, (E-H) *cer2-5* mutant Stages A-D, (I-L) *Gl2* transgene expressed in *cer2-5* mutant background Stages A-D. And, the adaxial surfaces of Arabidopsis flowers of genotypes: (M-P) *Gl2* transgene expressed in wild-type background Stages A-D, (Q-T) *CER2* transgene expressed in *cer2-5* mutant background Stages A-D



Supplemental Figure 4 Cont'd

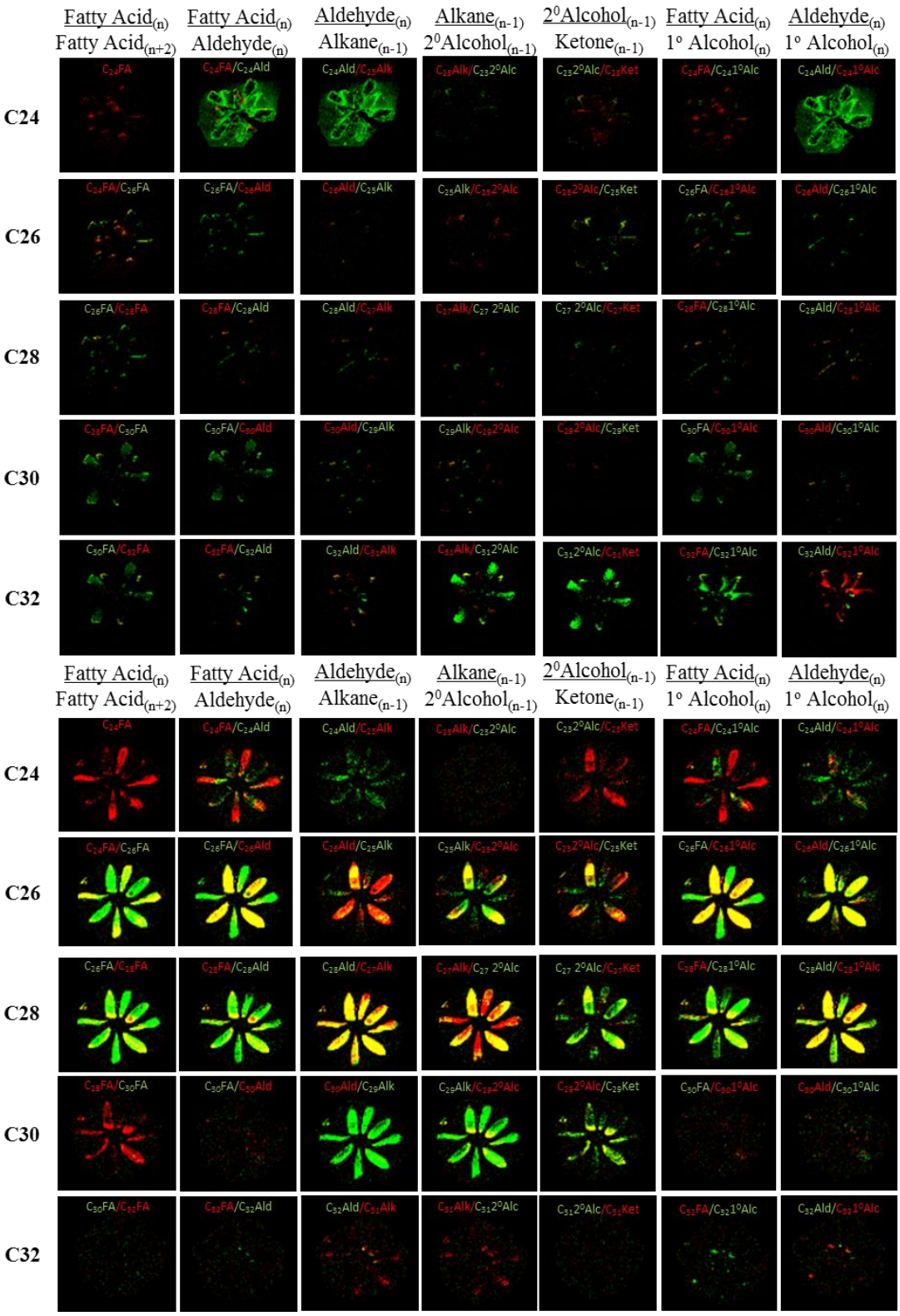


Supplemental Figure 4 Cont'd



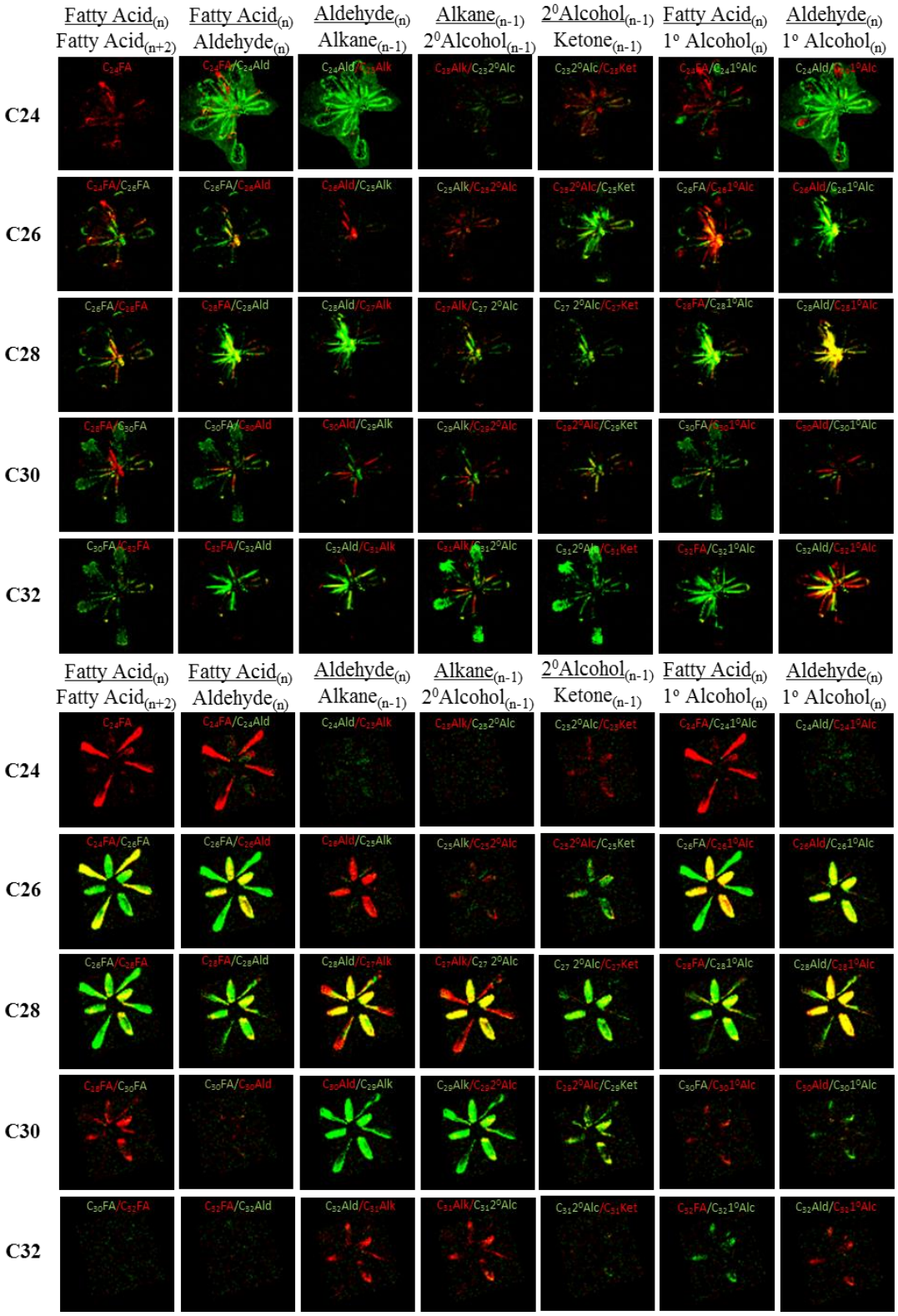
Supplemental Figure 4 Cont'd

(F)



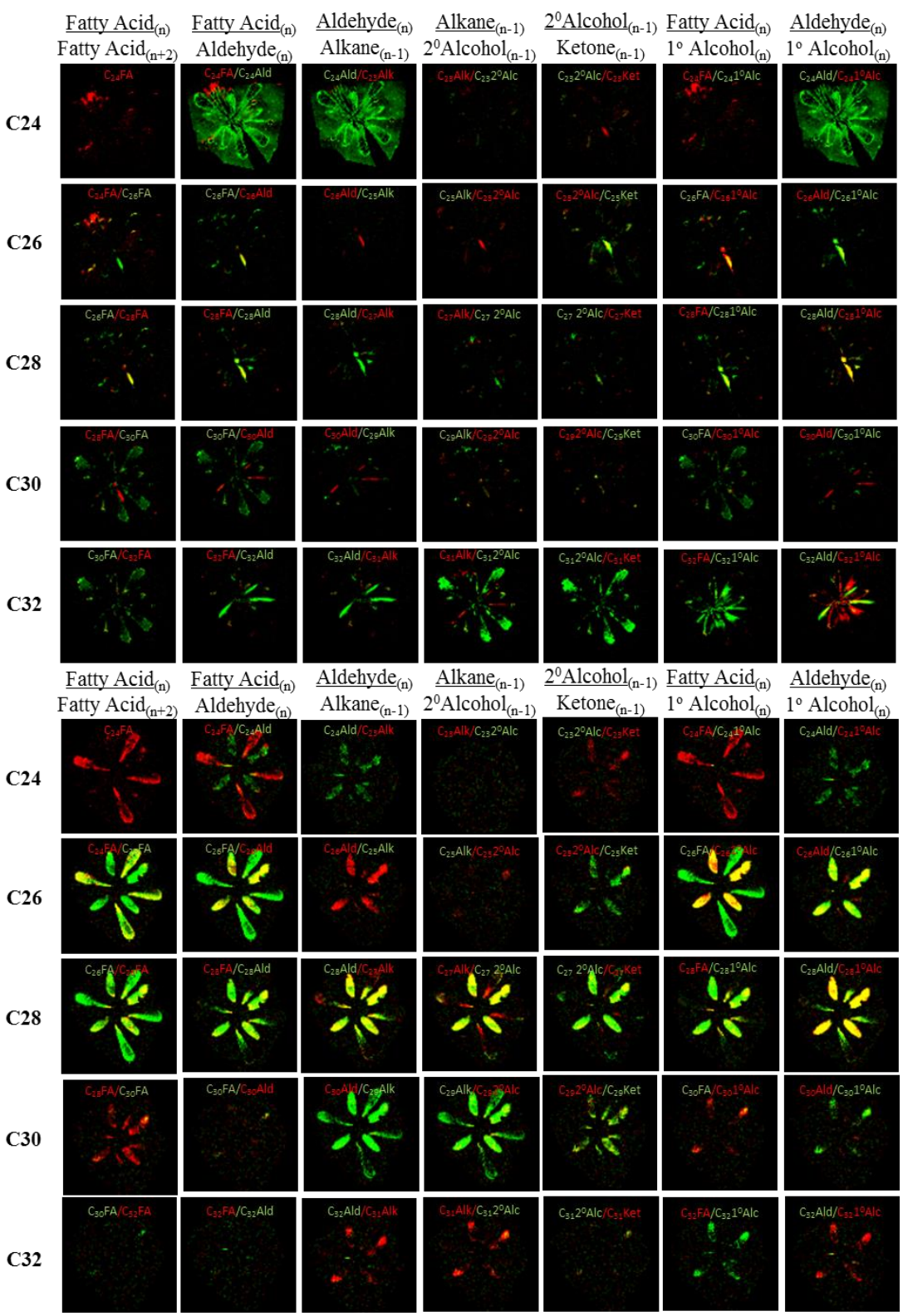
Supplemental Figure 4 Cont'd

(G)



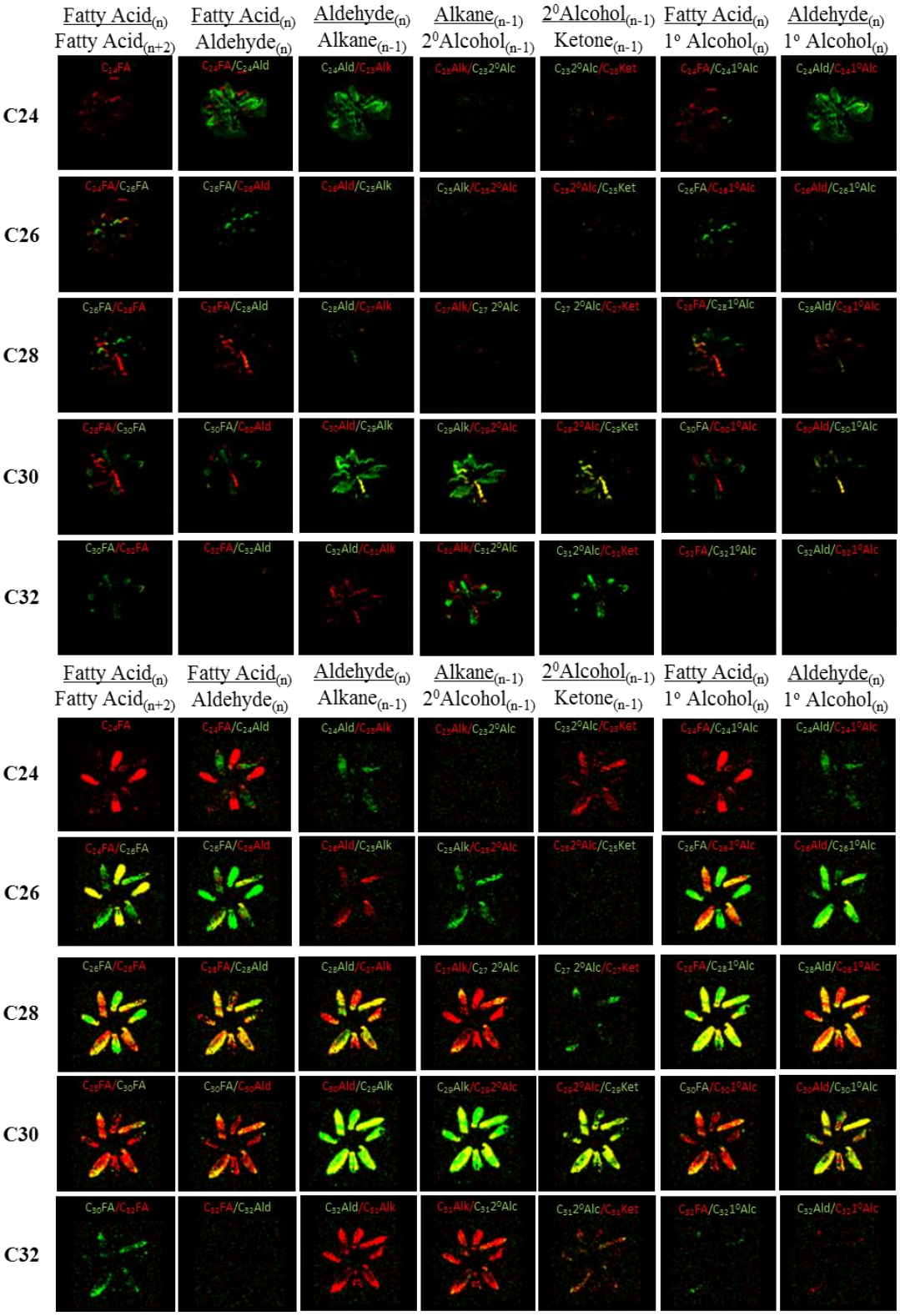
Supplemental Figure 4 Cont'd

(H)

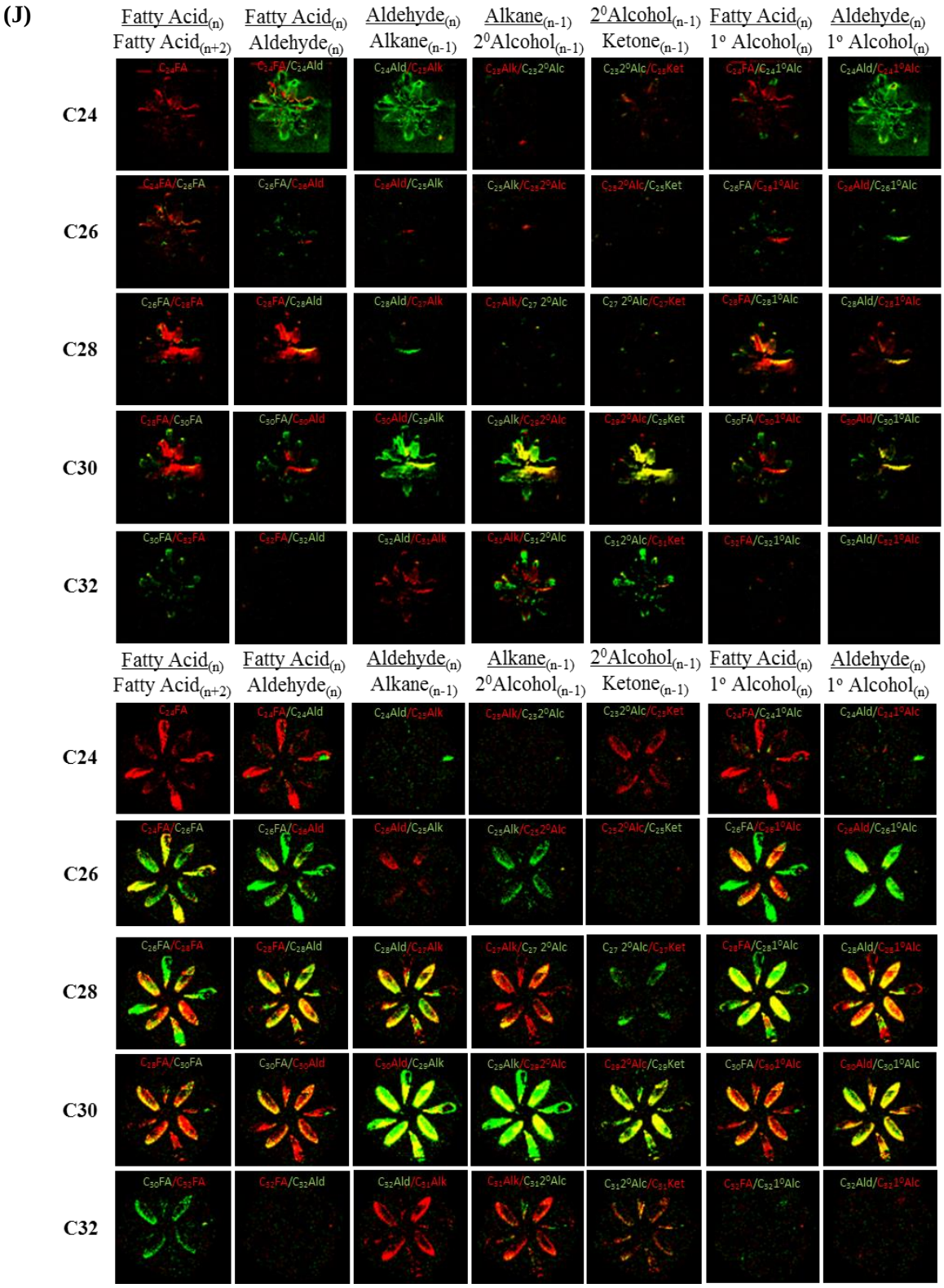


Supplemental Figure 4 Cont'd

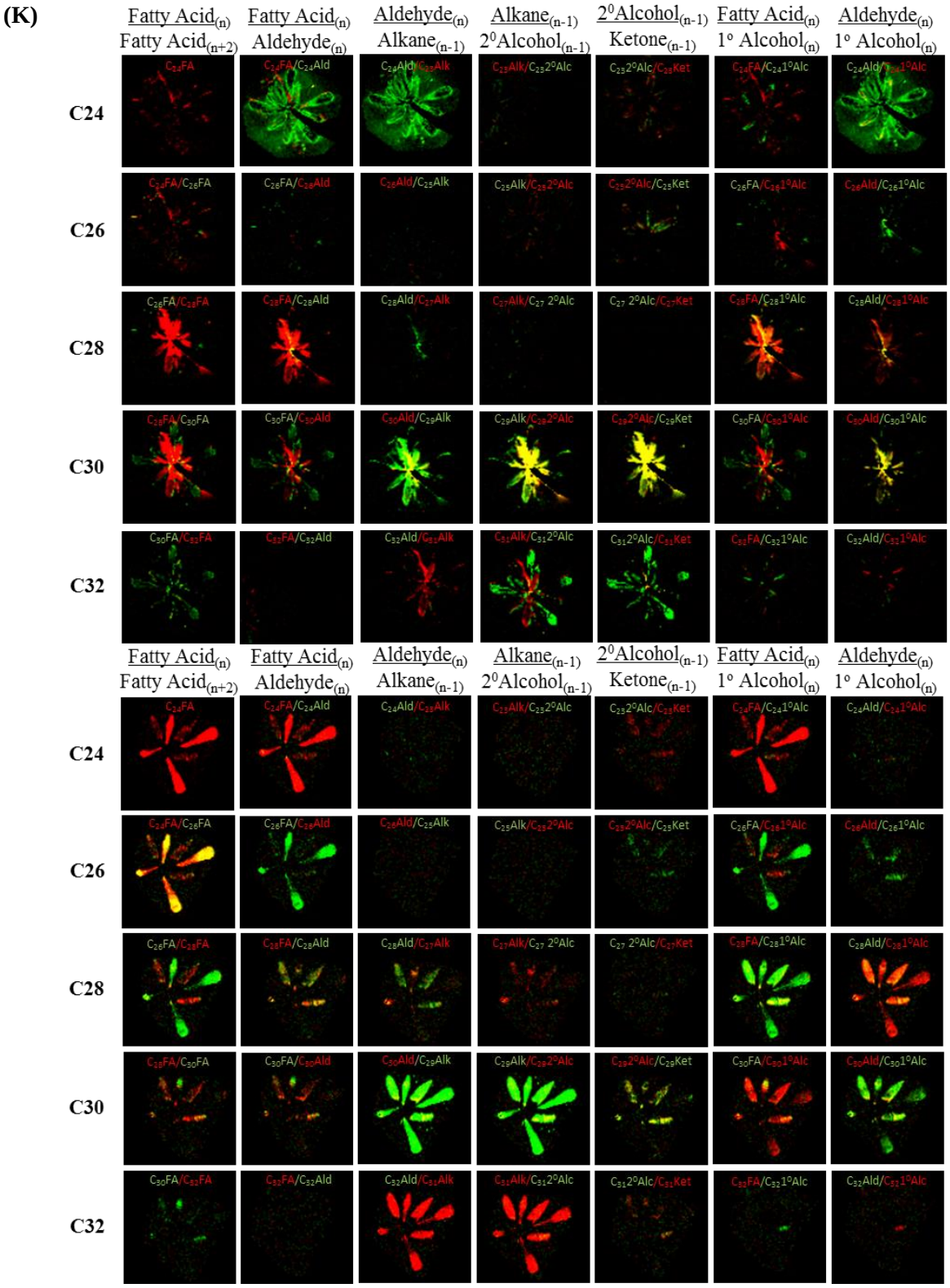
(I)



Supplemental Figure 4 Cont'd

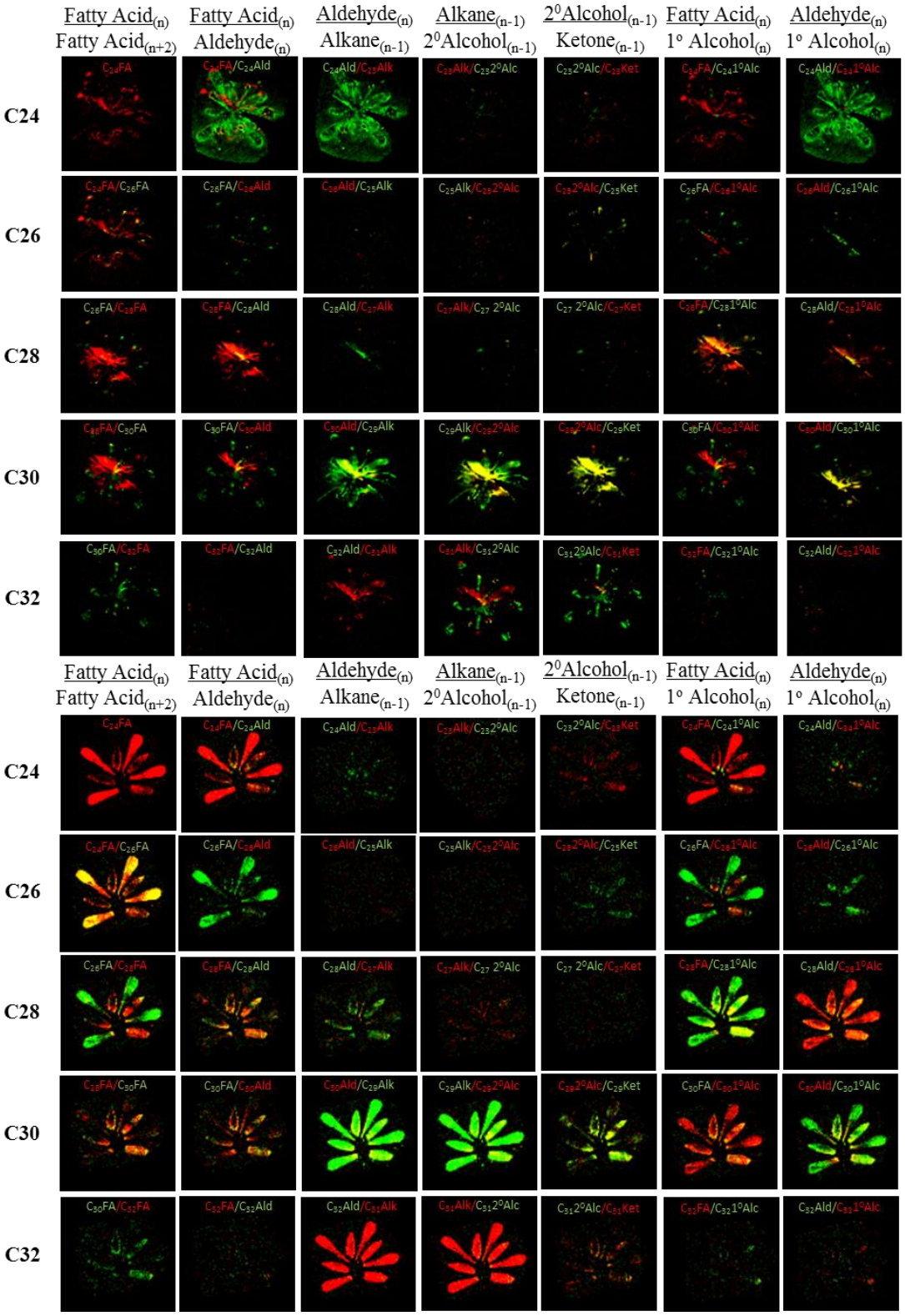


Supplemental Figure 4 Cont'd

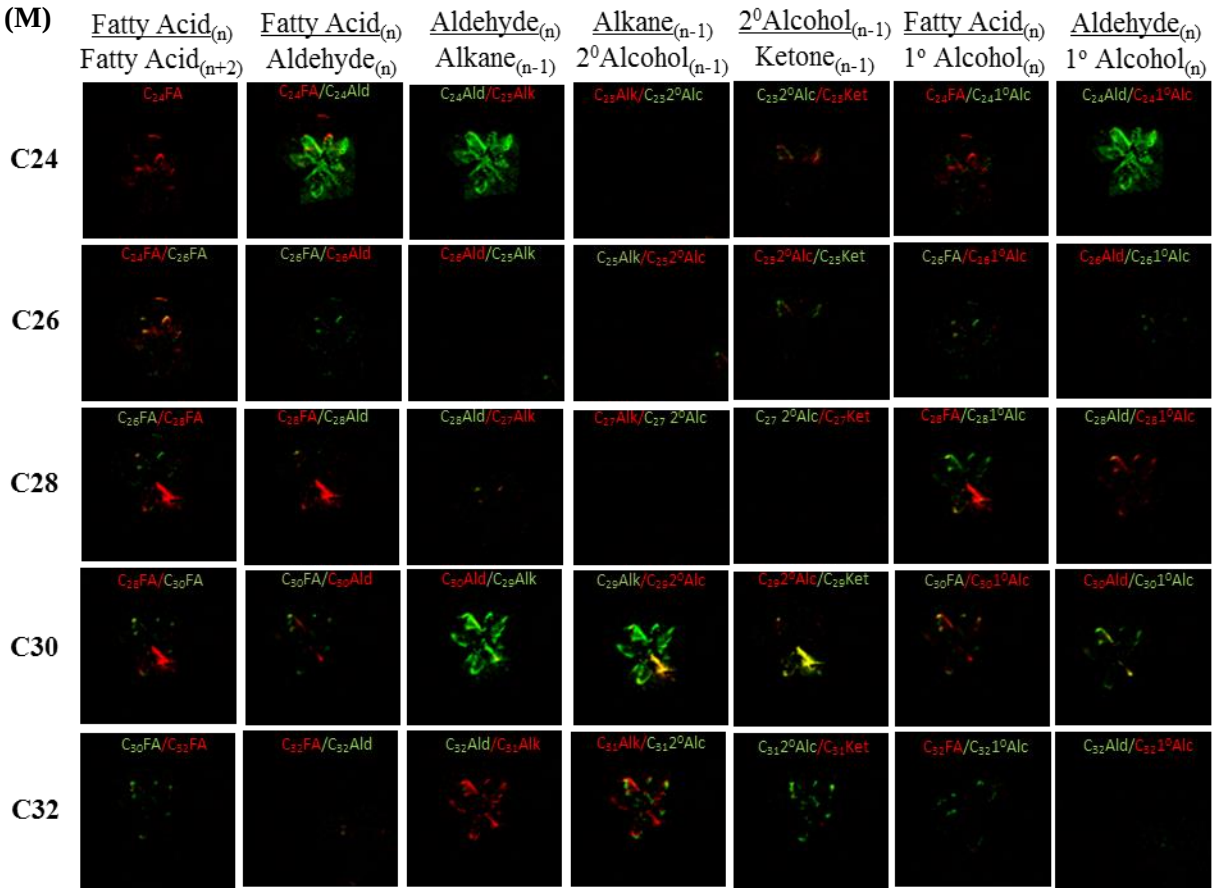


Supplemental Figure 4 Cont'd

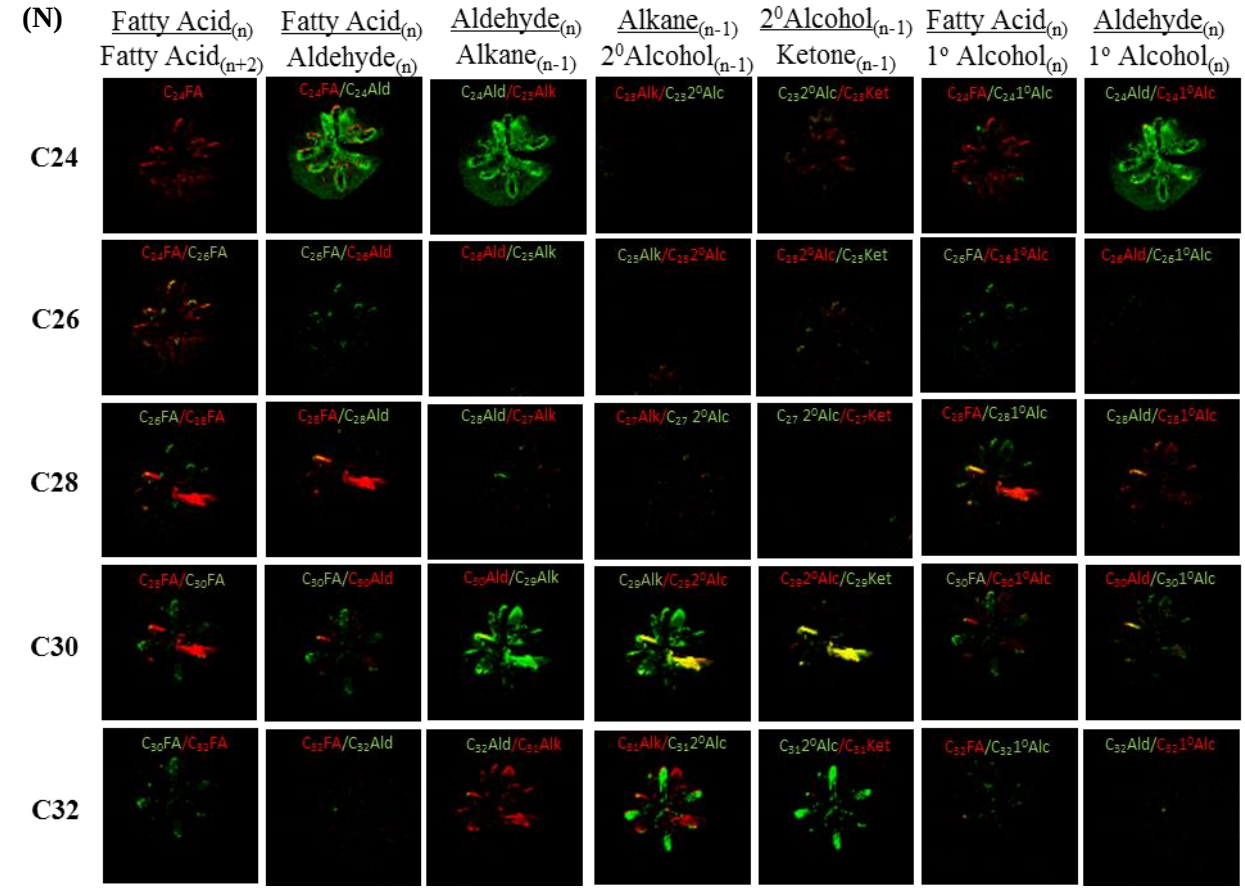
(L)



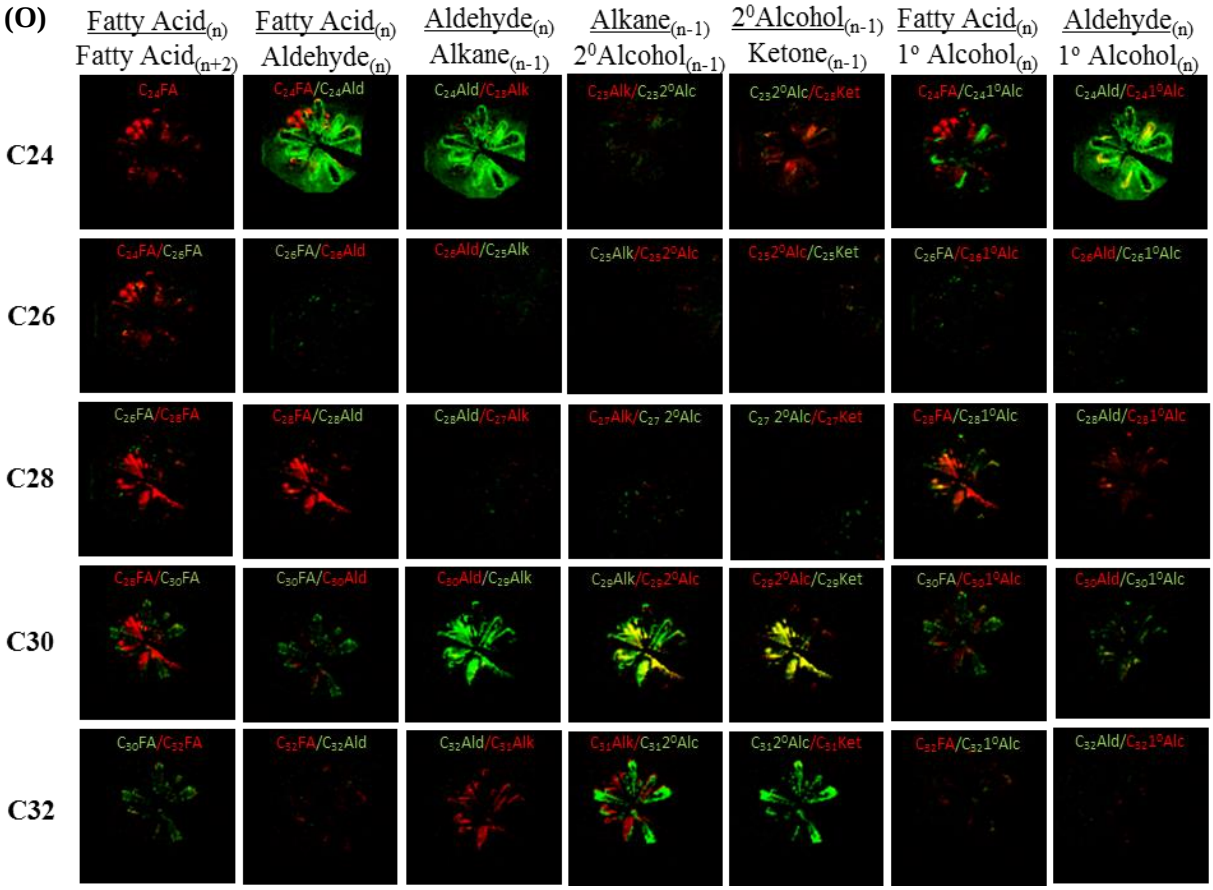
Supplemental Figure 4 Cont'd



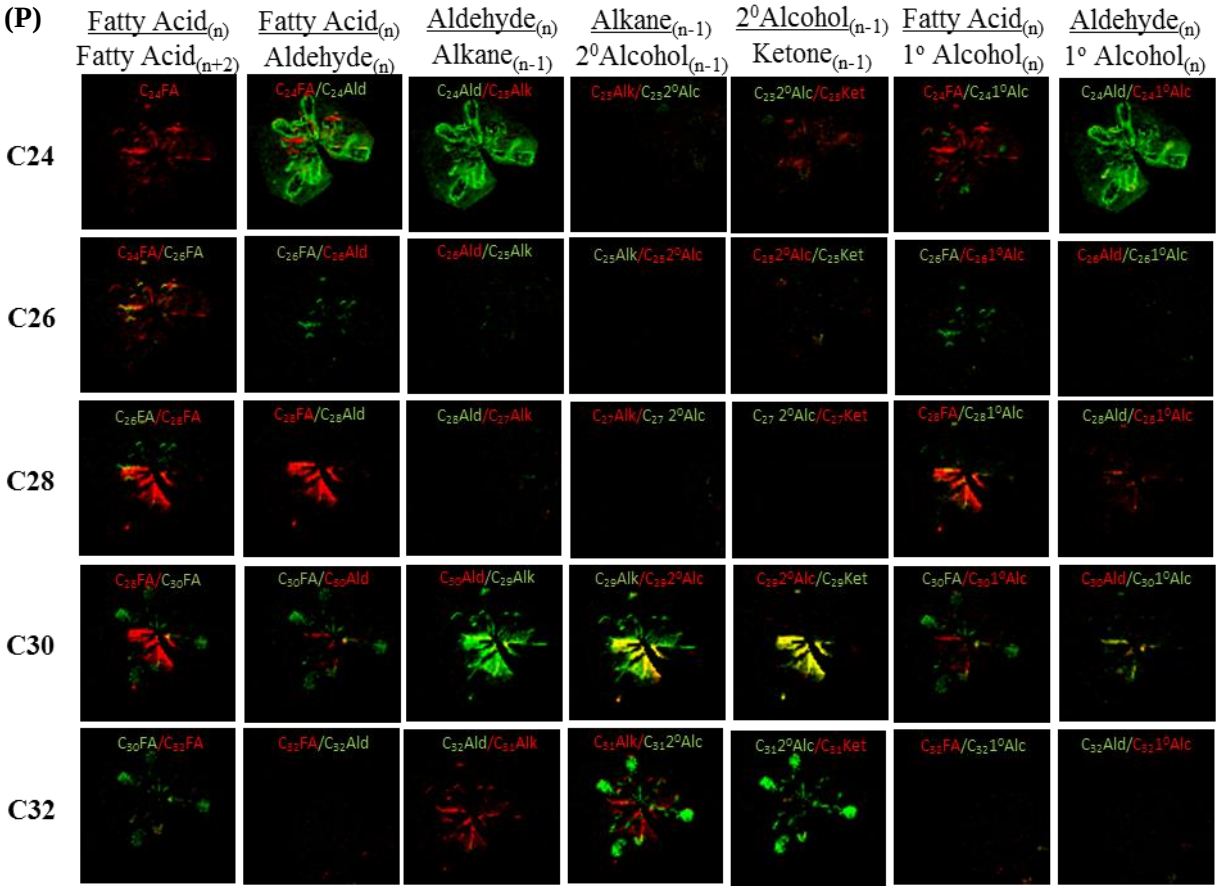
Supplemental Figure 4 Cont'd



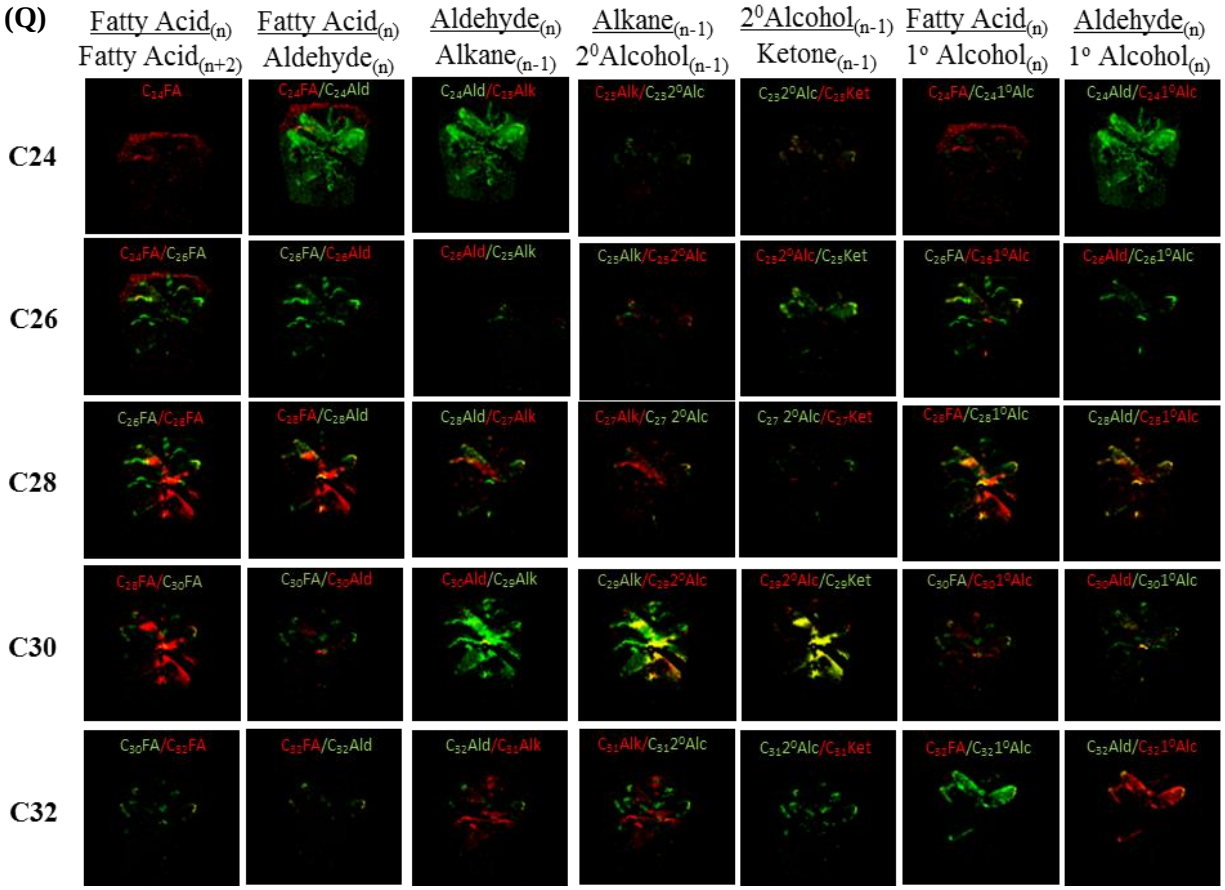
Supplemental Figure 4 Cont'd



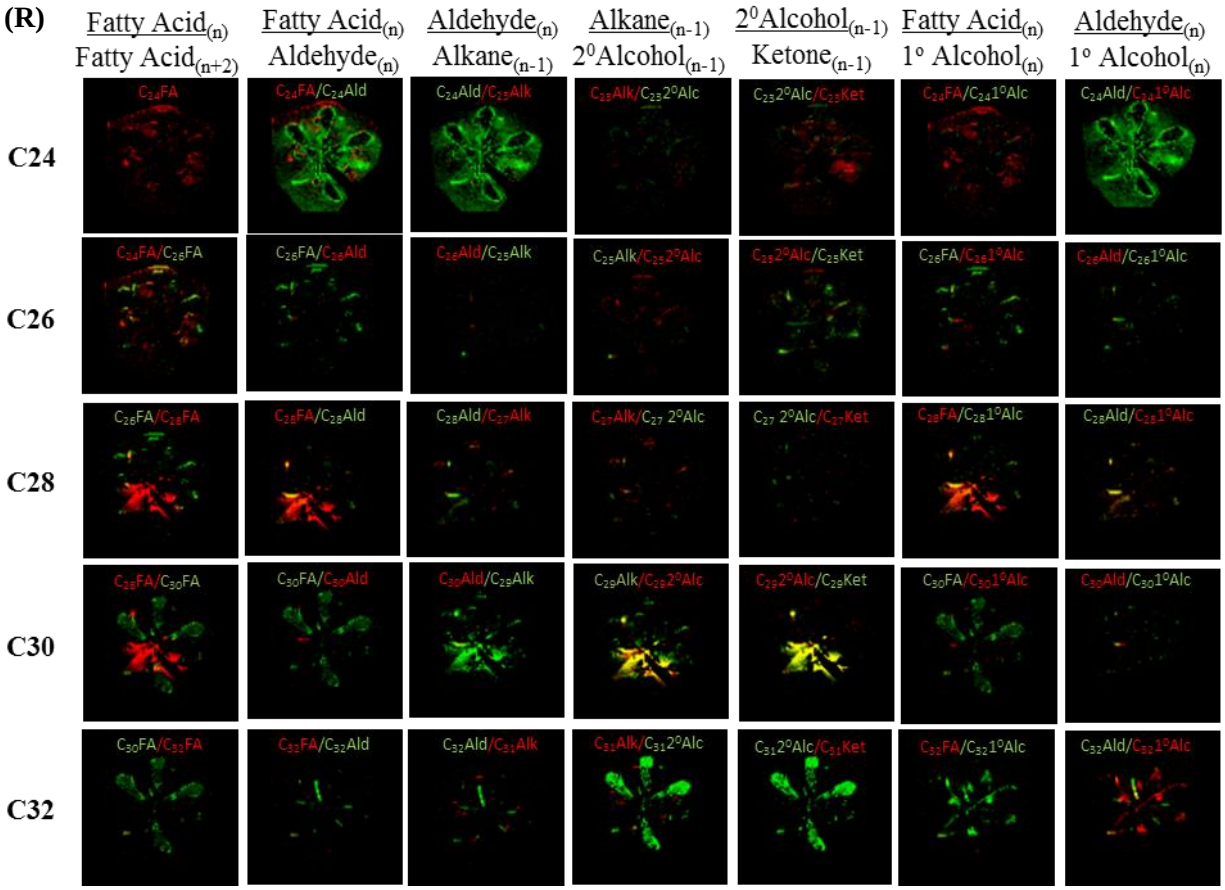
Supplemental Figure 4 Cont'd



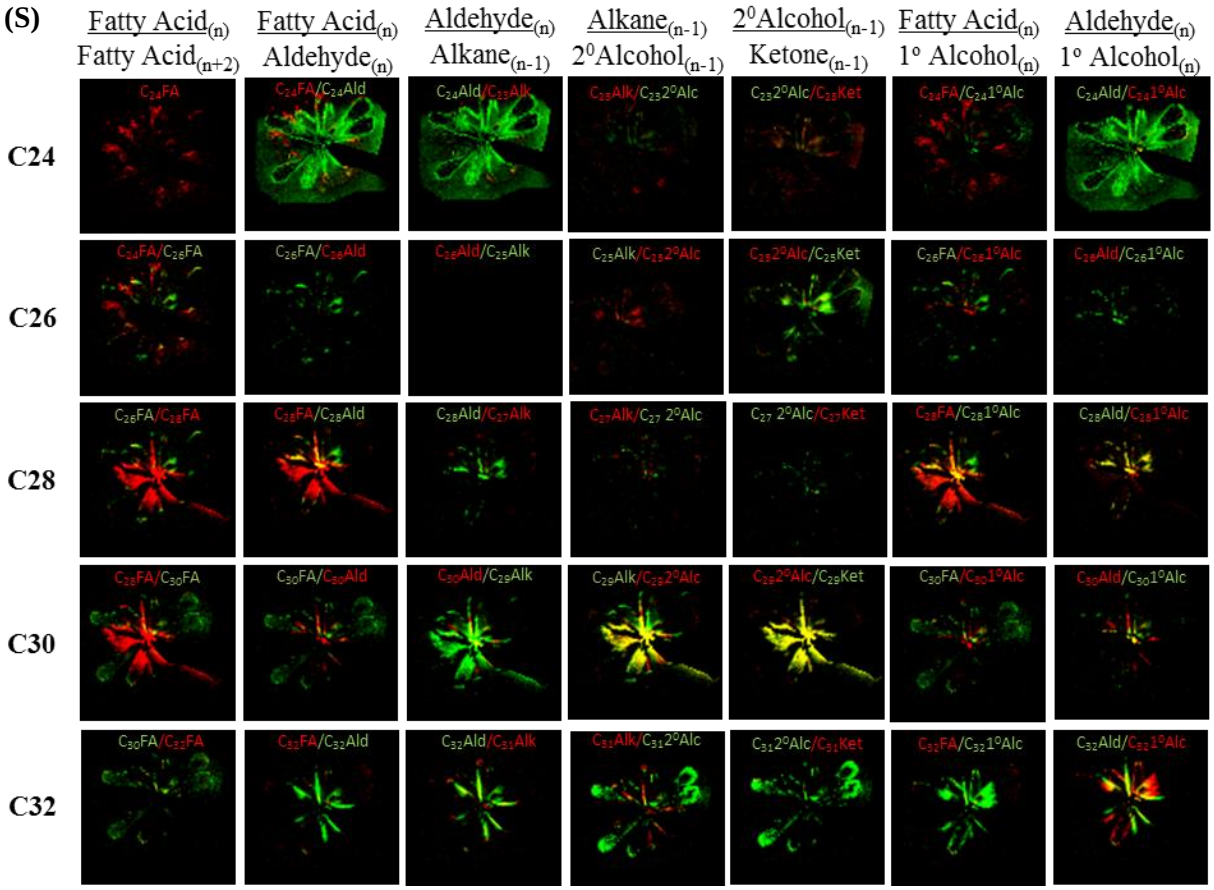
Supplemental Figure 4 Cont'd



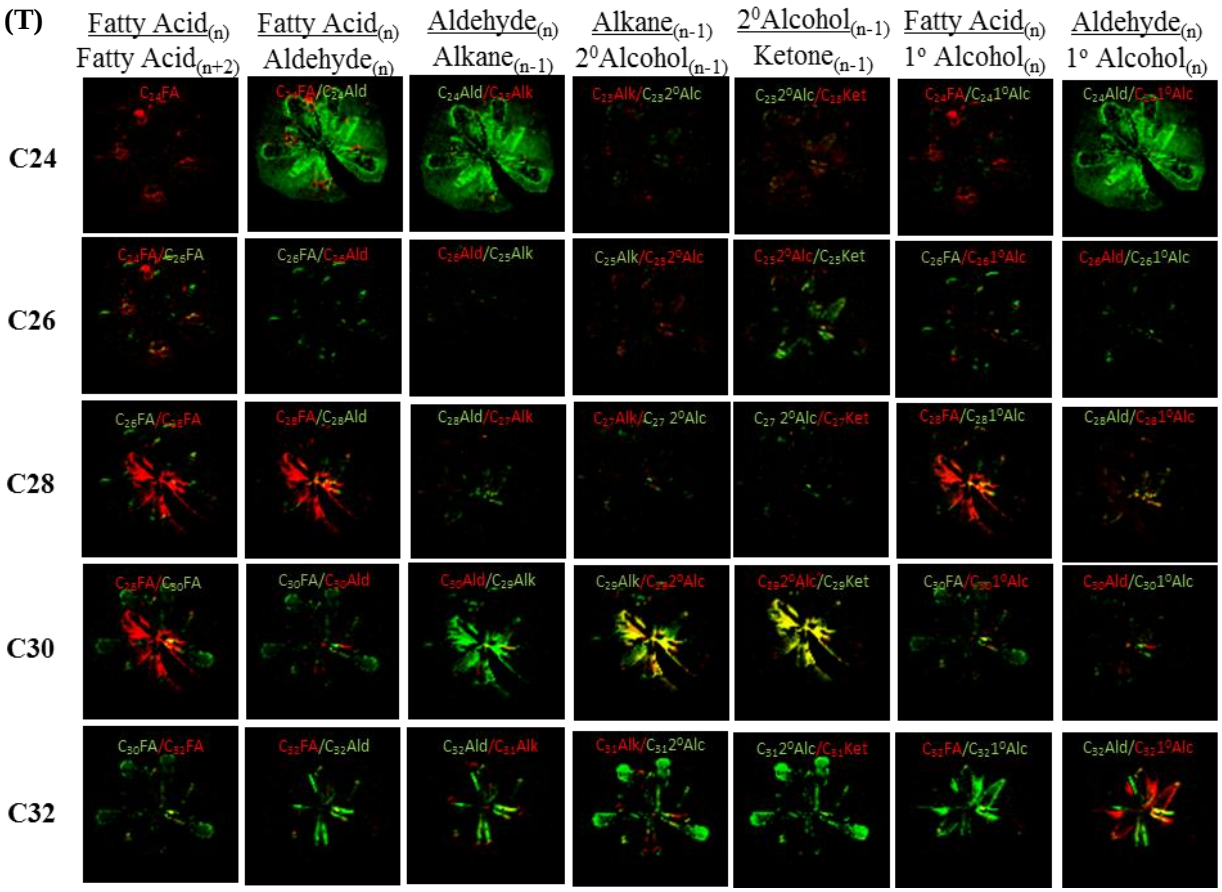
Supplemental Figure 4 Cont'd



Supplemental Figure 4 Cont'd



Supplemental Figure 4 Cont'd


(T)
Fatty Acid(n) Fatty Acid(n+2)
Fatty Acid(n) Aldehyde(n)
Aldehyde(n) Alkane(n-1)
Alkane(n-1) 2⁰Alcohol(n-1)
2⁰Alcohol(n-1) Ketone(n-1)
Fatty Acid(n) 1° Alcohol(n)
Aldehyde(n) 1° Alcohol(n)
C24
C26
C28
C30
C32

Supplemental Table 1. Flowering time of developing Arabidopsis flowers. These data were obtained from the indicated genotypes, which included the non-transgenic wild-type (Col-0) or *cer2-5* mutant, these two lines transgenically overexpressing the *Zea mays Glossy2 (ZmGl2)*, or the *Zea mays Glossy2-like (ZmGl2-like)* genes, and the *CER2* transgene in *cer2-5* mutant. Time of flowering at four stages of flower development are provided ± standard error (n= 2-4 replicates).

Time of Flowering (h)							
Genotype / Developmental Stage	WT (Col-0)	*cer2-5*	*cer2-5+CER2*	*cer2-5+ZmGl2*	WT+*ZmGl2*	*cer2-5+ZmGl2-like*	WT+*ZmGl2-like*
Stage A	00.00 ± 0.00	00.00 ± 0.00	00.00 ± 0.00	00.00 ± 0.00	00.00 ± 0.00	00.00 ± 0.00	00.00 ± 0.00
Stage B	02.09 ± 0.05	02.09 ± 0.09	02.09 ± 0.05	02.75 ± 0.48	02.09 ± 0.05	02.09 ± 0.05	02.67 ± 0.67
Stage C	09.50 ± 0.96	15.50 ± 0.50*	09.33 ± 0.67*	19.25 ± 3.09	12.04 ± 0.04*	15.00 ± 4.04	14.06 ± 0.06*
Stage D	33.50 ± 2.63	31.50 ± 4.50	22.00 ± 3.06	29.50 ± 3.77	24.25 ± 2.25*	25.00 ± 6.35	22.67 ± 0.67*
Asterisks black and red denote time points that are statistically significant from wild-type and *cer2-5* mutant respectively, as determined by Student's T-test (p-value <0.05)							

Supplemental Table 2. Biomass of developing Arabidopsis flowers. These data were obtained from the indicated genotypes, which included the non-transgenic wild-type (Col-0) or *cer2-5* mutant, these two lines transgenically overexpressing the *Zea mays Glossy2 (ZmGl2)*, or the *Zea mays Glossy2-like (ZmGl2-like)* genes, and the *CER2* transgene in *cer2-5* mutant. The biomass of flowers are estimated from the average of pooled (6-13) flowers at each stage of development performed, and estimated as the mean of 2-3 replicates ± standard deviation.

Weight of Arabidopsis single flowers (mg)							
Genotype / Developmental Stage	WT (Col-0)	*cer2-5*	*cer2-5+CER2*	*cer2-5+ZmGl2*	WT+*ZmGl2*	*cer2-5+ZmGl2-like*	WT+*ZmGl2-like*
Stage A	0.63 ± 0.037	0.56 ± 0.095	0.40 ± 0.089	0.41 ± 0.078	0.44 ± 0.042*	0.65 ± 0.050	0.69 ± 0.073
Stage B	0.73 ± 0.085	0.65 ± 0.059	0.47 ± 0.094	0.42 ± 0.028*	0.50 ± 0.056*	0.68 ± 0.021	0.71 ± 0.065
Stage C	0.84 ± 0.065	0.73 ± 0.128	0.59 ± 0.014	0.51 ± 0.016	0.72 ± 0.028	0.77 ± 0.069	0.82 ± 0.014
Stage D	0.94 ± 0.148	0.81 ± 0.097	0.62 ± 0.003	0.50 ± 0.049*	0.72 ± 0.021	0.95 ± 0.012	1.04 ± 0.186
Asterisks black and red denote the biomass of flowers that are statistically significant from wild-type and *cer2-5* mutant respectively, as determined by Student's T-test (p-value <0.05)							

Supplemental Table 3. Rate of extracellular cuticular lipid accumulation during development of Arabidopsis flowers. These data were obtained from the indicated genotypes, which included the non-transgenic wild-type (Col-0) or *cer2-5* mutant, and these two lines transgenically overexpressing the *Zea mays Glossy2-like (ZmGl2-like).* Data represents the cuticular lipid accumulation of initial stages (Stages A and B) and later stages (Stages C and D) of flower development per hour.

Metabolite Class	Alkyl Chain length	Wild-type (Col-0)		*cer2-5*		*ZmGl2-like* in *cer2-5*		*ZmGl2-like* in WT	
		Rate in A-B transition	Rate in C-D transition	Rate in A-B transition	Rate in C-D transition	Rate in A-B transition	Rate in C-D transition	Rate in A-B transition	Rate in C-D transition
		nmol h^{-1} g^{-1} Fresh Weight							
Fatty Acids	20	0.019	0.003	0.024	0.010	0.015	0.007	-0.011	0.008
	22	0.019	0.001	0.010	0.004	0.015	0.002	0.006	0.006
	24	0.106	-1E-04	0.130	-0.003	0.148	-0.014	0.024	-0.017
	26	0.044	0.002	0.133	0.015	0.040	-0.002	0.003	0.003
	28	-3E-05	-1E-06	0.174	0.015	0.040	-0.010	-4E-06	-9E-06
	30	-3E-05	0.001	-3E-05	-2E-06	-8E-06	3E-04	-4E-06	0.003
Primary Alcohol	24	0.012	0.001	0.006	0.002	0.030	-0.003	0.011	0.004
	26	0.109	0.014	0.478	0.116	0.388	0.071	0.057	0.021
	28	0.264	0.037	0.409	0.178	0.319	0.065	0.166	0.041
	30	0.112	0.017	1E-04	7E-05	0.054	0.012	0.111	0.007
Aldehyde	30	-3E-05	0.008	-3E-05	-2E-06	0.034	0.003	-0.021	0.011
Alkane	25	-0.021	-0.003	-0.046	-0.014	-0.017	-0.024	-0.006	-0.025
	27	0.976	-0.026	2.148	-0.113	0.696	-0.117	0.348	-0.112
	29	17.513	0.923	1.978	-0.247	8.784	0.700	12.067	0.923
	31	0.422	0.011	0.214	-0.014	0.375	0.023	0.845	0.013
Sec. Alcohol	27	-3E-05	-1E-06	0.025	0.007	0.009	0.003	-4E-06	-9E-06
	29	1.730	0.230	0.003	2E-04	0.518	0.228	1.219	0.390
Ketone	29	1.223	0.313	0.014	0.002	0.455	0.340	0.875	0.481

Supplemental Table 4. Rate of extracellular cuticular lipid accumulation during development of Arabidopsis flowers. These data were obtained from the indicated genotypes, which included the non-transgenic wild-type (Col-0) or *cer2-5* mutant, and these two lines transgenically overexpressing the *Zea mays Glossy2 (ZmGl2)*, and the CER2 transgene in cer2-5 mutant. Data represents the cuticular lipid accumulation of initial stages (Stages A and B) and latter stages (Stages C and D) of flower development per hour.

Metabolite Class	Akyl Chain length	Wild-type (Col-0)		*cer2-5*		*ZmGl2* in *cer2-5*		*ZmGl2* in WT		*CER2* in *cer2-5*	
		Rate in A-B transition	Rate in C-D transition	Rate in A-B transition	Rate in C-D transition	Rate in A-B transition	Rate in C-D transition	Rate in A-B transition	Rate in C-D transition	Rate in A-B transition	Rate in C-D transition
		nmol h^{-1} g^{-1} Fresh Weight									
Fatty Acids	20	0.161	0.004	-0.059	-0.002	0.005	0.019	-0.003	0.024	0.618	-0.015
	24	0.590	0.014	0.405	-0.054	0.581	0.063	0.725	0.092	1.155	-0.055
	26	0.185	0.003	0.231	-0.029	0.162	0.018	0.039	0.018	0.743	-0.082
Primary Alcohol	22	-0.021	-0.029	0.013	0.009	0.190	0.273	-0.005	0.097	0.386	-0.002
	26	0.003	0.037	0.597	0.594	0.071	0.432	0.144	0.047	0.195	0.212
	28	0.106	0.089	1.935	0.663	0.173	0.459	0.247	0.091	0.546	0.689
	30	0.027	0.018	-0.004	0.004	0.107	0.236	0.152	0.044	0.271	0.075
Alkane	25	0.306	-0.012	-0.246	-0.041	-0.106	-0.078	0.063	-0.013	0.174	0.035
	27	0.598	-0.037	2.168	-0.005	0.387	-0.044	0.466	-0.097	0.969	0.032
	29	11.882	1.322	2.509	0.097	13.793	5.592	24.136	2.186	4.442	1.944
	31	1.684	0.155	-0.014	-0.035	3.964	1.838	12.689	0.926	1.163	0.019
Sec. Alcohol	29	0.982	0.488	0.024	0.008	1.377	1.597	2.729	0.577	0.753	0.551
Ketone	29	1.319	0.505	0.011	0.010	1.289	1.954	3.213	0.957	1.753	0.638

CHAPTER 7: CHARACTERIZATION OF THE *GLOSSY2* GENE FAMILY BY HETEROLOGOUS EXPRESSION IN YEAST

Manuscript in preparation to be submitted to PLoSOne

[1]Liza Esther Alexander[a,b], Katelyn R. Campbell[c], Kenna E. Stenback[a,b], Kayla Flyckt[a,b], Michael A. Schelling[a], Ludmila Rizhsky[a,b], Basil J. Nikolau[a,b].

Abstract

In planta studies have demonstrated that GLOSSY2 and GLOSSY2-LIKE may be involved in the elongation of very long chain fatty acids (VLCFAs) from 26:0 or 28:0 to 32:0. However the mechanism by which this occurs is unclear. VLCFAs are synthesized by a fatty acyl-CoA elongase complex (FAE) that iteratively catalyzes a series of 4 reactions: a) Claisen condensation; b) 3-ketoacyl reduction; c) 3-hydroxyacyl dehydration; and d) enoyl reduction. The first of these reactions is catalyzed by either 3-ketoacyl-CoA synthase (KCS) or ELO (Elongation Defective-Like) enzymes, and the subsequent 3 reactions are catalyzed by 3-ketoacyl-CoA reductase (KCR), 3-hydroxyacyl-CoA dehydratase (HCD), and enoyl-CoA reductase (ECR). The use of *Saccharomyces cerevisiae* as a host for heterologous expression has proven advantageous in reconstituting the plant FAE complex. Using this system, we evaluated whether the co-expression of either GLOSSY2 or GLOSSY2-LIKE with individual maize FAE enzymatic components affects the VLCFA product profile of the FAE system. Conducting such co-expression

[a] Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA, USA.
[b] Center for Metabolic Biology, Iowa State University, Ames, IA, USA.
[c] Departments of Genetics, Development, and Cell Biology and Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA, USA.

experiments in diploid yeast strains indicate that both GLOSSY2 and GLOSSY2-LIKE could affect the VLCFA product profile, however such results could not be recapitulated when they were conducted in yeast haploid strains. These results suggest that the apparent regulatory role of GLOSSY2 or GLOSSY2-LIKE to affect FAE product profile may be an artifact of adding and mixing plant components to the endogenous FAE enzyme complex.

Introduction

The maize *Glossy2* and *Glossy2-like* genes are functional homologs of the Arabidopsis *CER2* gene (Alexander, 2019; this thesis). Mutations in the *glossy2* or *cer2* genes affect the normal accumulation of epicuticular lipids, and both these mutations appear to affect the ability to elongate a specific chain-length of fatty acids; in maize the *gl2* mutation blocks the ability to elongate beyond C30 chain length (1), whereas in Arabidopsis mutations in the *cer2* gene appears to affect the ability to elongate fatty acids beyond C26 or C28 chain lengths to C30 fatty acids (2,3). Recent studies using heterologous expression in yeast indicate that CER2-type proteins interact with the fatty acid elongation system to affect the fatty acid product profile of the enzyme complex (4–6). However, the exact biochemical mechanism of how this occurs, and how this extrapolates to the *in planta* effects of mutations in these genes is still unclear.

Fatty acid elongase (FAE) complexes carry out the elongation of preexisting fatty acids to acyl chains of 20 carbons and longer. This complex has four enzymatic components, the first of which catalyzes the condensation reaction between a precursor acyl-CoA and malonyl-CoA. This component was initially identified as the product of the *fae1* locus of Arabidopsis (7), which encodes a 3-ketoacyl-CoA synthase (KCS), and subsequent studies have identified that up to 21 KCS homologs occur in Arabidopsis (8–10) and 27 homologs in maize (11). Characterizations of the yeast FAE system identified an alternative enzyme that catalyzes this initial condensation

reaction, elongase-type (ELO) enzyme (12). The ELO type condensing enzymes occur in animals, fungi, and plants, and KCS enzymes occur exclusively in plants (13–15).

In addition to the enzyme(s) that catalyze the initial condensation reaction of FAE, 3 other enzymes constitute the FAE complex: 1) 3-ketoacyl-CoA reductase (KCR), 2) enoyl-CoA reductase (ECR), and 3) 3-hydroxyacyl-CoA dehydratase (HCD). In maize, genetic redundancy mainly occurs at the first two reactions of FAE; in addition to the 27 putative KCS enzymes, the maize genome encodes 6 putative ELO enzymes, and 2 active KCR enzymes (16). In contrast single copy genes appear to encode the HCD and ECR component enzymes.

The maize KCR, HCD and ECR components have been functionally characterized by heterologous expression in yeast (11,16). These studies have established an *in vivo* yeast-based system for functionally reconstituting the maize FAE complex (11; Stenback et al., unpublished). The simplicity of the yeast reconstituted FAE system makes it possible to mix and match individual KCS, ELO and KCR maize components, and thus study the role of apparent gene redundancy among these components (11). In this study, we used this reconstituted FAE system to investigate if maize GLOSSY2 and GLOSSY2-LIKE proteins interact with maize FAE components, and thus modify the *in vivo* generated very long chain fatty acid (VLCFA) product profile.

Results and Discussion

Computational identification of maize KCS homologs

Recent studies in yeast have identified 27 KCS maize homologs (11). Yeast expression studies with the Arabidopsis CER2 suggest that it acts by interacting with a specific FAE KCS component enzyme (also identified as *CER6*), and facilitating a specific iteration of the FAE elongation cycle (4,5). We explored whether this model is applicable to the maize GLOSSY2 and GLOSSY2-LIKE proteins. Initially we used the Aramemnon database (17) to explore the

phylogenetic diversity of maize KCS isozymes (11) and identified 27 KCS candidates. Figure 1 illustrates that these 27 homologs categorize 11 clades (α to λ), Clade α/β includes two maize KCS genes, *ZmKCS4* (GRMZM2G393897) and *ZmKCS15* (GRMZM2G164974) and Clade γ includes the two maize KCS candidates that belong to the *CER6*-containing clade, *ZmKCS5* (AC233893.1_FGT003) and *ZmKCS6* (GRMZM2G164974) (11). The latter two maize KCS homologs (*ZmKCS5* or *ZmKCS6*) share ~90% sequence identity with the AtCER6 protein family (Supplemental Figure 1) and were selected for further characterization.

Effect of co-expressing maize *Glossy2* or *Glossy2-like* with maize KCS homologs in wild-type diploid yeast strains

By co-expressing *Gl2* or *Gl2-like* with *ZmKCS5* or *ZmKCS6* in yeast cells, we evaluated whether there are any interactions that could affect VLCFA product profiles of the resulting yeast strains. These co-expression experiments were conducted in two diploid strains, INVSc1 and BY4743 (18,19). Figure 2 shows the VLCFA profiles engendered by the co-expression of the maize genes, under the control of the galactose-inducible, *GAL1* promoter from an episomal plasmid-based vector. Both of these strains normally yield similar VLCFA titers (~7.0 µmol/g DW), ranging from 20:0 to 28:0, and the most abundant VLCFA is 26:0 (Figure 2). Co-expressing *Gl2* or *Gl2-like* with either *ZmKCS5* or *ZmKCS6* in the INVSc1 strain induced an approximately 2-fold increase in the VLCFA titer (Figures 2C and 2D), and this is specifically associated with increased titers of 24:0, 26:0 and 28:0 (Figures 2A and 2B). The fact that these increased titers occur only upon co-expression of the KCS and GL2 homologs, and not when these 4 maize genes are individually expressed (Figure 2) are indicative of interactions between these proteins. These observations are therefore consistent with those reported with the Arabidopsis CER2 and the 3

CER2-LIKE proteins (4,5) that affect increased accumulation of specific VLCFAs (i.e., 26:0 and 28:0).

However, *in planta* characterizations of genetic manipulations of either the *Gl2* or *Gl2-like* genes indicate that these genes may be involved in producing VLCFAs beyond 26:0/28:0, and up to 32:0 (1; Alexander, 2019; this thesis), a trait that is not recapitulated in these yeast expression studies. Therefore, we considered whether this may be associated with: 1) the type of promoter that was used in the co-expression experiments; or 2) by the particular yeast strain that was used (INVSc1), which is a commercial product, whose provenance is unknown. Hence, both the *Gl2* and *Gl2-like* genes were individually co-expressed with either *ZmKCS5* or *Zm*KCS6 under the control of the constitutive, glycerol-3-phosphate dehydrogenase (*GPD*) promoter (20). Furthermore, these latter experiments where conducted a commonly used laboratory diploid strain, BY4743 with a well-defined pedigree (18, 19). Figure 3 compares the results of these latter experiments.

In contrast to the results obtained with the use of the *GAL1* promoter in the INVSc1 strain, the use of the *GPD* promoter induced additional VLCFAs (30:0) that do not occur in the parental INVSc1 strain (Figures 3A and 3B), and this occurred without any significant changes in the abundance of the precursor fatty acids (20:0 to 28:0). Additionally, this novel capability was dependent upon the co-expression of either *Gl2* or *Gl2-like* with either of the two *ZmKCS* isozymes.

This outcome was not recapitulated when co-expression was conducted in the BY4743 diploid strain (Figures 3C and 3D). Specifically, the co-expression of *Gl2* or *Gl2-like* with *ZmKCS5* did not change the VLCFA profile from the parental wild-type strain or from that of the individually expressed control strains. Although co-expression of *ZmKCS6* with *Gl2-like* produced

detectable quantities of 30:0, this was not consistent among all replicates, and thus was not statistically significant. However, the co-expression of *ZmKCS6* with *Gl2* produced statistically significant amounts of 30:0. These distinct observations between the two diploid strains and the different outcomes produced by the use of the *GPD* (20) and *GAL1* promoters (21), indicates that additional insights are required to dissect the complexity associated with the potential genetic anomalies associated with mixing FAE components from heterologous systems.

Effect of co-expressing maize *Glossy2* or *Glossy2-like* with maize KCS homologs as replacements of yeast FAE components

There are multiple potential factors that can contribute to the complexity of co-expressing the maize KCS and GL2 homologs in yeast diploid strains. One of these is the fact that the maize KCS and GL2 homologs are being genetically added to an intact endogenous FAE system, and this endogenous yeast FAE system does not utilize a KCS-type condensing-enzymes, but rather utilizes a combination of three ELO-type condensing-enzymes (ScELO1, ScELO2 and ScELO3) (12,14). This complexity was overcome by taking advantage of the efficacy of yeast genetics, and thereby individually co-express maize KCS and GL2 pairs in a strain, which partially lacks the endogenous yeast FAE system; namely, the maize KCS and GL2 pairs were systematically co-expressed in haploid *elo* mutants.

Specifically, we co-expressed 4 KCS isozymes (the CER6-type *ZmKCS5* and *ZmKCS6*, and *ZmKCS4* and *ZmKCS15*), with either *ZmGl2* or *ZmGl2-like* in haploid strains (i.e., BY4742 and derivatives). These parental haploid strains carried either a single *elo3* mutant allele or a combination of *elo2 elo3* double mutation. The *elo3* single mutant strain is viable, and served as the control strain for interpreting the results from the *elo2 elo3* double mutant strain which is normally a synthetic lethal, but can be rescued by the expression of a maize *ZmKCS* gene (11).

Strains individually expressing *ZmKCS5, ZmKCS6, ZmGl2* or *ZmGl2-like* in the *elo3* single mutant strain, and those co-expressing *ZmKCS5* or *ZmKCS6* with either *ZmGl2* or *ZmGl2-like* do not alter the VLCFA product profiles of the strains (Figures 4A and 4B). Whereas the strains individually expressing *ZmKCS4* or *ZmKCS15* were able to restore the *elo3* mutation by the production of 26:0 although not to wild-type levels (Figures 4C and 4D). More significantly, the former co-expressing strains did not induce the accumulation of the novel VLCFA product, 30:0, which was observed in the diploid strains (Figure 3). As an additional control, *AtCER2* was co-expressed with either one of the 4 ZmKCS isozymes (i.e, *ZmKCS5, ZmKCS6, ZmKCS4, ZmKCS15*), however the co-expression also failed to alter the VLCFA product profiles (Figure 4). These results suggest, that the apparent interactions between the ZmKCS homologs and GL2 or GL2-LIKE to affect the production of longer chain VLCFAs which was observed in the diploid strains may be associated with the complexity associated with "mixing" endogenous FAE catalytic components (specifically ScELO3) with "foreign" plant components that do not normally cross such genetic/evolutionary barriers.

Additional experiments further explored these complexities by using the *elo2 elo3* double mutant haploid strain (generated by sporulating the diploid BY4743 (11). This synthetically lethal haploid strain, can be rescued by the episomal expression of a ZmKCS homolog (11), and in this study we explored how coexpressing a ZmKCS with either GL2 or GL2-LIKE affected the FAE product profile (Figure 5). These strains provide an additional platform for evaluating potential KCS-GL2 interactions in the absence of both endogenous ScELO2 and ScELO3 components. However, to aid in the interpretation of the outcomes from these genetic manipulations an additional control strain was developed by the episomal expression of the *ScELO3* gene in the *elo2 elo3* double mutant strain. The *elo2 elo3* double mutant strain, rescued by the ectopic expression

of *ScELO3* generates a VLCFA profile with chain lengths of up to 26:0, and 22:0 being the most abundant VLCFA.

In subsequent experiments we replaced the episome that is expressing the *ScELO3*-rescuing gene with the 4 ZmKCS homologs (i.e., *ZmKCS4, ZmKCS5, ZmKCS6 or ZmKCS15*) (11) and each of these successfully rescued the lethality associated with the *elo2 elo3* double mutant. These 4 ZmKCS-rescued strains produce VLCFAs that are up to 28:0, with small quantities of 30:0; these are at levels similar to those observed with the diploid strains, INVSc1 and BY4743 (Figure 5). More significantly, when either *Gl2* or *Gl2-like* is co-expressed with any of the 4 ZmKCS homologs, no additional changes in the VLCFA profiles were generated. These observations therefore demonstrate that any of the 4 maize KCSs are capable of producing 30:0, which is reflective of their function in maize, but their co-expression with either *Gl2 or Gl2-like* does not further alter the VLCFA profiles of the resulting strains (Figure 5). Therefore these observations are not consistent with the effects we observed when these co-expression experiments were conducted in the INVSc1 diploid strain, which could be interpreted as indicating interactions between GL2 homologs and KCS isozymes. Such interactions have been inferred for the *CER2-KCS* interactions, which were conducted in the INVSc1 diploid strain (5).

Effect of co-expressing *Glossy2* and *Glossy2-like* with other FAE components

The FAE system consists of a set of potentially redundant enzyme-components that iteratively catalyze a series of 4 reactions that collectively elongate a preexisting fatty acyl-CoA by 2-carbons per cycle. The *in planta* chemotype of the *gl2* mutant indicates that this gene product affects the terminal 1 or 2 iterations of the FAE cycle, i.e., the elongation of 28:0 to 30:0 and 32:0 (Alexander, 2019; this thesis). A potential model to adjudicate these datasets is to infer that the *Gl2*-gene product is a regulator of the terminal two elongation cycles. This could be mediated by

specific interactions between GL2 and any one of the FAE components that catalyze the terminal 2 iterations of the FAE complex. We therefore evaluated whether *Gl2* or *Gl2-like* interact with other (non-KCS) FAE components to affect the product profile of the system, these being specifically *ZmELO1*, *ZmKCR1*, *ZmKCR2*, and *ZmHCD*.

ZmELO1 (GRMZM2G037152) is one of 6 ELO homologs that occurs in the maize genome, and it has been shown to be functional in complementing the lethality associated with the yeast *elo2 elo3* double mutant (11). Expressing *ZmELO1* either individually or in combination with either *Gl2* or *Gl2-like* in a wild-type or *elo2 elo3* double mutant generated near identical VLCFA profiles (Figure 6). We conclude that GL2 and GL2-LIKE proteins do not interact with ZmELO1 and are not required for supporting the ZmELO1 dependent elongation process.

Prior molecular genetic characterizations of the *GLOSSY8a* (ZmKCR1) and *GLOSSY8b* (ZmKCR2) genes established their role in cuticular lipid deposition as enzymatic components of the FAE system (16,22,23,11). In the current study we evaluated if ZmKCR1 or ZmKCR2 maybe interacting with either *Gl2* are *Gl2-like* and thereby affect changes in VLCFA profiles produced by the FAE system. Specifically, either *ZmKCR1* or *ZmKCR2* was co-expressed with either *Gl2* or *Gl2-like* in yeast strains that lacked the endogenous KCR (i.e., a yeast strain carrying a *kcr* mutant allele). The yeast *kcr* mutant strain is lethal, but can be rescued by the expression of either *ZmKCR1* or *ZmKCR2* (11). Comparing the VLCFA profiles of the resulting rescued yeast strains indicate that the expression of either *ZmKCR1* or *ZmKCR2* produced VLCFA profiles that are similar to those that occur in the wild-type strain, with 26:0 being the most abundant (Figures 7 and 8). The co-expression of either *Gl2* or *Gl2-like* with these maize homologs does induce quantitative changes in the VLCFA profiles, but these are relatively minor (ranging between 20%

and 80% increase), but there is no qualitative change in this profile, as was observed when *Gl2* or *Gl2-like* were expressed in diploid strains that resulted in the appearance of 30:0 (Figure 3).

The third reaction of the FAE elongation cycle is catalyzed by HCD and recent studies have demonstrated that the maize HCD enzyme has broad substrate capabilities and can rescue the lethality associated with a yeast *hcd* mutation (11). Therefore, we investigated if *Gl2* or *Gl2-like* interacts with ZmHCD (GRMZM2G151087) by co-expressing the two combinations in the yeast *hcd* mutant strain. Figure 9 shows the comparison of the VLCFA profiles of these *ZmHCD*-rescued *hcd* strains with and without the co-expression of either *Gl2* or *Gl2-like.* Although the VLCFA profile of the yeast strain is not altered when *ZmHCD* is expressed in the wild-type background, this profile is strikingly different when *ZmHCD* is expressed in the *hcd* mutant background. Specifically, there are large and statistically significant increases in the accumulation of 3-hydroxy-VLCFAs and VLCFAs of between 20 and 24 carbon chain lengths, and these changes are concomitant with decreased accumulation of 3-hydroxy-VLCFAs and VLCFAs of longer chain lengths. However, the co-expression of either *Gl2* or *Gl2-like* with *ZmHCD* did not further alter the VLCFA profiles of the rescued strains (Figure 9).

Summary

Understanding the complexity of the fatty acid elongation process is complicated by genetic redundancy and the fact that the stoichiometry of the components in the elongation complex is unknown. *In planta* studies on *Glossy2* and *Glossy2-like* demonstrate that these genes influence the fatty acid elongation between 26:0/28:0 and 32:0 (Alexander, 2019; this thesis). Considering that fatty acid elongation is carried out by 4 enzymes of the FAE complex, we evaluated whether *Gl2* and *Gl2-like* interacts with individual components of the FAE complex by co-expression in yeast. Our findings demonstrate that GL2 or GL2-LIKE do not interact with

individual components of the FAE complex. This finding deviates from proposed interaction of CER2 homologs with KCS isozymes that have been proposed for the Arabidopsis and rice FAE systems (5,6). This is demonstrated by two key findings 1) four maize KCS orthologs (KCS4, KCS5, KCS6, or KCS15) can genetically complement the lethality associated with the yeast *elo2 elo3* double mutant strain, and these rescued strains can produce VLCFAs up to 30:0; and 2) the co-expression of GLOSSY2 or GLOSSY2-LIKE with any of the four maize KCS orthologs does not further alter the yeast VLCFA profile. Furthermore, we also demonstrate that neither GLOSSY2 nor GLOSSY2-LIKE affect changes in VLCFA profiles when they are co-expressed in yeast with other maize FAE components (i.e., ELO1, KCR1, KCR2 or HCD).

Although the co-expression of KCS and *Gl2* homologs in diploid yeast strains are capable of producing novel VLCFA, these findings could not be recapitulated in haploid strains devoid of native elongation proteins i.e., ELO2 and ELO3. These observations raise a concern on how protein-protein interactions are deduced in yeast, especially when mixing plant and yeast components. For example, recent characterizations of Arabidopsis KCS isozymes demonstrate that when KCS16, KCS6 and CER2-LIKE proteins are co-expressed in yeast the strains are capable of producing novel fatty acids, up to 38:0 fatty acids (24). However, unlike the extensive studies done on KCS in Arabidopsis, the substrate specificities of maize KCS orthologs are yet to be explored. Our study has reported the characterization of 4 maize KCS orthologs, and 23 additional KCS isozymes are yet to be characterized. Therefore, whether other KCS homologs are required for a potential interaction with *Gl2* homologs is still to be determined.

Future studies that utilize yeast FAE replacement strategies that express all maize FAE components would be one way to address if the VLCFA lipid profile changes observed in plants

expressing either *Glossy2* or *Glossy2-like* is due to an interaction with the FAE complex or independent of such interactions i.e., associated with other proteins unrelated to the FAE complex.

Materials and Methods

Molecular Cloning

The ORFs of *Glossy2* (GRMZM2G098239) was codon-optimized for yeast using GeneOptimizer (GeneArt, LifeTechnologies)), *Glossy2-like* (GRMZM2G315767) was codon optimized for Arabidopsis using OptimumGene^{TM} (GenScript, Piscataway, NJ; www.genscprit.com), and the coding region of the CER2 full-length cDNA (AT4G24510) was amplified with Platinum pfx DNA polymerase (Invitrogen) and the amplified fragment was purified using the QIAquick gel extraction kit (Qiagen). These genes were cloned into high-copy episomal plasmids, pAG426 (URA3), or pAG423 (HIS3) (Invitrogen, Carlsbad, CA) using either Gateway® cloning system (Invitrogen, Carlsbad, CA) or In-Fusion® cloning (Takara Bio USA, Inc., Mountain View, CA). Depending on the experiment, the episomal plasmids were under the control of galactose-inducible promoter (GAL1) or constitutive glyceraldehyde-3-P-dehydrogenase (GPD) promoter. All recombinant yeast shuttle vectors were confirmed by DNA sequencing, and were maintained in *E.coli* TOP10 cells (Invitrogen, Carlsbad, CA) using Luria Bertani (LB) media supplemented with the appropriate antibiotics.

Yeast strains and media

The yeast strains, INVSc1, BY4743, BY4742, and yeast mutant *elo3* were maintained in YPD (yeast peptone dextrose) or synthetic complete (SC) dropout media. Yeast strains carrying mutations in the endogenous FAE component genes and complemented by maize FAE components were provided by Campbell et al. and details of these strains can be found in Campbell et al., 2019. These strains and all other yeast strains carrying maize *Glossy2* and *Glossy2-like* were

selected by their ability to grow on minimal medium (SD) lacking the appropriate amino acid or nucleobase (e.g., uracil). The yeast strains that contained episomal plasmids with GAL1 promoter were induced with the inclusion of 2% galactose in SC medium. Yeast cultures were grown according to standard procedures in appropriate media at 30^{o}C (25,26). For fatty acid analysis experiments all strains were grown for 72-hours, except for the INVSc1+GAL1 experiment which was grown for 48-h.

Yeast Transformation

Plasmids were transformed into yeast using a standard lithium acetate transformation protocol (27). Briefly, 3 µL of salmon sperm DNA, 1 µg of plasmid DNA, and 100 µL transformation mix (200 µL of 2 M Lithium Acetate; 800 µL of 50% PEG; 7.7 µL of β-mercaptoethanol) were added to a 1.5 mL microcentrifuge tube and vortexed. A large yeast colony is added to the mix, vortexed, and incubated on a shaker at 37°C for 30 min. Samples are then centrifuged at 3000 rpm for 5 min, the supernatant is discarded and residual pellet is resuspended in 200 µL of water and plated on selective solid media. The plates are incubated at 30°C for 2-3 days. The empty vector was transformed into each strain and used as a control.

Fatty Acid Analysis

FAMEs were synthesized by transmethylation of lyophilized cells at 80°C for 1 h using 1mL 5% sulfuric acid in methanol and 10 µg of nonadecanoic acid as an internal standard. After cooling, 2 mL of 0.9% aqueous NaCl was added for phase separation, and fatty acyl chains were extracted thrice with 1 mL of 4:1 hexane:chloroform. Samples were dried under nitrogen gas to a final volume of 250 µL. One microliter of the samples were analyzed by GC-MS or GC-FID.

Samples that were analyzed by GC-MS were silylated with 50 µL *N*,*O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS) (Sigma-

Aldrich), prior to injection. GC-FID analysis was conducted with an Agilent 6890 GC, equipped with a DB-1 MS capillary column (15 m x 0.25 mm x 0.25 μm, Agilent 122-0112). Chromatography was conducted with helium gas, at a flow-rate of 1.2 mL/min, and an inlet temperature at 280ºC. The column oven temperature was initially held at 80 °C for 1 min, then ramped at 15 °C/min to 230°C and held for 2 min and, then ramped at 15°C/min to 340°C and held there for 2 min. For peak identification purposes chromatograms were compared with FAME standards (8:0-30:0) and parallel GC-MS analysis.

Acknowledgements

The authors acknowledge Dr. Alexis A. Campbell for the maize FAE component yeast replacement strains, and Drs. Ann Perera and Lucas Showman of the WM Keck Metabolomics Research Laboratory (Iowa State University, Ames, IA) for assistance in metabolic analyses. This work was partly funded by grants from NSF, and by the State of Iowa through the Center for Metabolic Biology.

References

1. Bianchi. Glossy mutants of maize. VI. Chemical constituents of glossy-2 epicuticular waxes. Maydica [Internet]. 1975 Jan 1;20(4):165–73. Available from: https://eurekamag.com/research/000/389/000389438.php
2. McNevin, J.P., Woodward, W., Hannoufa, A., Feldmann, K.A., and Lemieux B. Isolation and characterization of eceriferum (cer) mutants induced by T-DNA insertions in Arabidopsis thaliana. 1991;(Mori 1982).
3. Jenks M a., Tuttle H a., Eigenbrode SD, Feldmann K a. Leaf Epicuticular Waxes of the Eceriferum Mutants in Arabidopsis. Plant Physiol. 1995;108(1):369–77.
4. Haslam TM, Mañas-Fernández A, Zhao L, Kunst L. Arabidopsis ECERIFERUM2 is a component of the fatty acid elongation machinery required for fatty acid extension to exceptional lengths. Plant Physiol [Internet]. 2012;160(3):1164–74. Available from: http://www.plantphysiol.org/content/160/3/1164.short
5. Haslam TM, Haslam R, Thoraval D, Pascal S, Delude C, Domergue F, et al. ECERIFERUM2-LIKE Proteins Have Unique Biochemical and Physiological Functions in Very-Long-Chain Fatty Acid Elongation. Plant Physiol [Internet]. 2015;167(3):682–92.

Available from:
http://www.plantphysiol.org/content/early/2015/01/16/pp.114.253195.abstract?papetoc

6. Wang X, Guan Y, Zhang D, Dong X, Tian L, Qu LQ. A β-Ketoacyl-CoA Synthase Is Involved in Rice Leaf Cuticular Wax Synthesis and Requires a CER2-LIKE Protein as a Cofactor. Plant Physiol [Internet]. 2017 Feb 1;173(2):944 LP-955. Available from: http://www.plantphysiol.org/content/173/2/944.abstract

7. James DW, Lim E, Keller J, Plooy I, Ralston E, Dooner HK. Directed tagging of the Arabidopsis FATTY ACID ELONGATION1 (FAE1) gene with the maize transposon activator. Plant Cell [Internet]. 1995 Mar 1;7(3):309 LP-319. Available from: http://www.plantcell.org/content/7/3/309.abstract

8. Bessoule J-J, Lessire R, Cassagne C. Partial purification of the Acyl-CoA elongase of Allium porrum leaves. Arch Biochem Biophys. 1989;268(2):475–84.

9. Fehling E, Lessire R, Cassagne C, Mukherjee KD. Solubilization and partial purification of constituents of acyl-CoA elongase from Lunaria annua. Biochim Biophys Acta (BBA)/Lipids Lipid Metab. 1992;1126(1):88–94.

10. Joubès J, Raffaele S, Bourdenx B, Garcia C, Laroche-Traineau J, Moreau P, et al. The VLCFA elongase gene family in Arabidopsis thaliana : phylogenetic analysis, 3D modelling and expression profiling. An Int J Mol Biol Mol Genet Biochem. Dordrecht; 2008;67(5):547–66.

11. Campbell AA, Stenback KE, Flyckt K, Hoang T, Perera MADN, Nikolau BJ. A single-cell platform for reconstituting and characterizing fatty acid elongase component enzymes. Submitted to PLoSOne 2019

12. Toke DA, Martin CE. Isolation and characterization of a gene affecting fatty acid elongation in Saccharomyces cerevisiae. Isol Charact a gene Affect Fat acid elongation Saccharomyces cerevisiae. 1996;(31):18413–22.

13. Cinti DL, Cook L, Nagi MN, Suneja SK. The fatty acid chain elongation system of mammalian endoplasmic reticulum. Prog Lipid Res. 1992;31(1):1–51.

14. Oh CS, Toke DA, Mandala S, Martin CE. ELO2 and ELO3, homologues of the Saccharomyces cerevisiae ELO1 gene, function in fatty acid elongation and are required for sphingolipid formation. J Biol Chem. 1997;272(28):17376.

15. Tvrdik P, Westerberg R, Silve S, Asadi A, Jakobsson A, Cannon B, et al. Role of a New Mammalian Gene Family in the Biosynthesis of Very Long Chain Fatty Acids and Sphingolipids. J Cell Biol. 2000;149(3):707.

16. Dietrich CR, Perera MADN, D Yandeau-Nelson M, Meeley RB, Nikolau BJ, Schnable PS. Characterization of two GL8 paralogs reveals that the 3-ketoacyl reductase component of fatty acid elongase is essential for maize (Zea mays L.) development. Plant J. 2005 Jun;42(6):844–61.

17. Schwacke R, Schneider A, van der Graaff E, Fischer K, Catoni E, Desimone M, et al. ARAMEMNON, a Novel Database for Arabidopsis Integral Membrane Proteins. Plant Physiol [Internet]. 2003 Jan 1;131(1):16 LP-26. Available from: http://www.plantphysiol.org/content/131/1/16.abstract

18. Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, et al. Designer deletion strains derived from Saccharomyces cerevisiae S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast. 1998;14(2):115–32.

19. Winzeler EA, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, et al. Functional Characterization of the S. cerevisiae Genome by Gene deletion and Parallel Analysis. Science (80-). 1999;285.

20. Bitter GA, Egan KM. Expression of heterologous genes in Saccharomyces cerevisiae from vectors utilizing the glyceraldehyde-3-phosphate dehydrogenase gene promoter. Gene [Internet]. Elsevier; 1984 Dec 1 [cited 2019 Jan 12];32(3):263–74. Available from: https://www.sciencedirect.com/science/article/pii/0378111984900027

21. West Jr RW, Yocum RR, Ptashne M. Saccharomyces cerevisiae GAL1-GAL10 divergent promoter region: location and function of the upstream activating sequence UASG. Mol Cell Biol [Internet]. 1984 Nov;4(11):2467–78. Available from: https://www.ncbi.nlm.nih.gov/pubmed/6392852

22. Xu X, Dietrich CR, Delledonne M, Xia Y, Wen TJ, Robertson DS, et al. Sequence analysis of the cloned glossy8 gene of maize suggests that it may code for a beta-ketoacyl reductase required for the biosynthesis of cuticular waxes. Plant Physiol. 1997;115(2):501–10.

23. Xu X, Dietrich CR, Lessire R, Nikolau BJ, Schnable PS. The Endoplasmic reticulum-associated maize GL8 protein is a component of the acyl-coenzyme A elongase involved in the production of cuticular waxes. Plant Physiol [Internet]. 2002;128(3):924–34. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=152205&tool=pmcentrez&rendertype=abstract

24. Hegebarth D, Buschhaus C, Joubès J, Thoraval D, Bird D, Jetter R. Arabidopsis ketoacyl-CoA synthase 16 (KCS16) forms C36/C38 acyl precursors for leaf trichome and pavement surface wax. Plant Cell Environ. 2017;40(9):1761–76.

25. Sherman F, Fink G, Hicks J. Methods in Yeast Genetics. 1986.

26. Rose MD, Winston F, Hieter P. Methods in Yeast Genetics, A Laboratory Course Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory; 1990.

27. Gietz RD, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 2002;350:87.

Figures

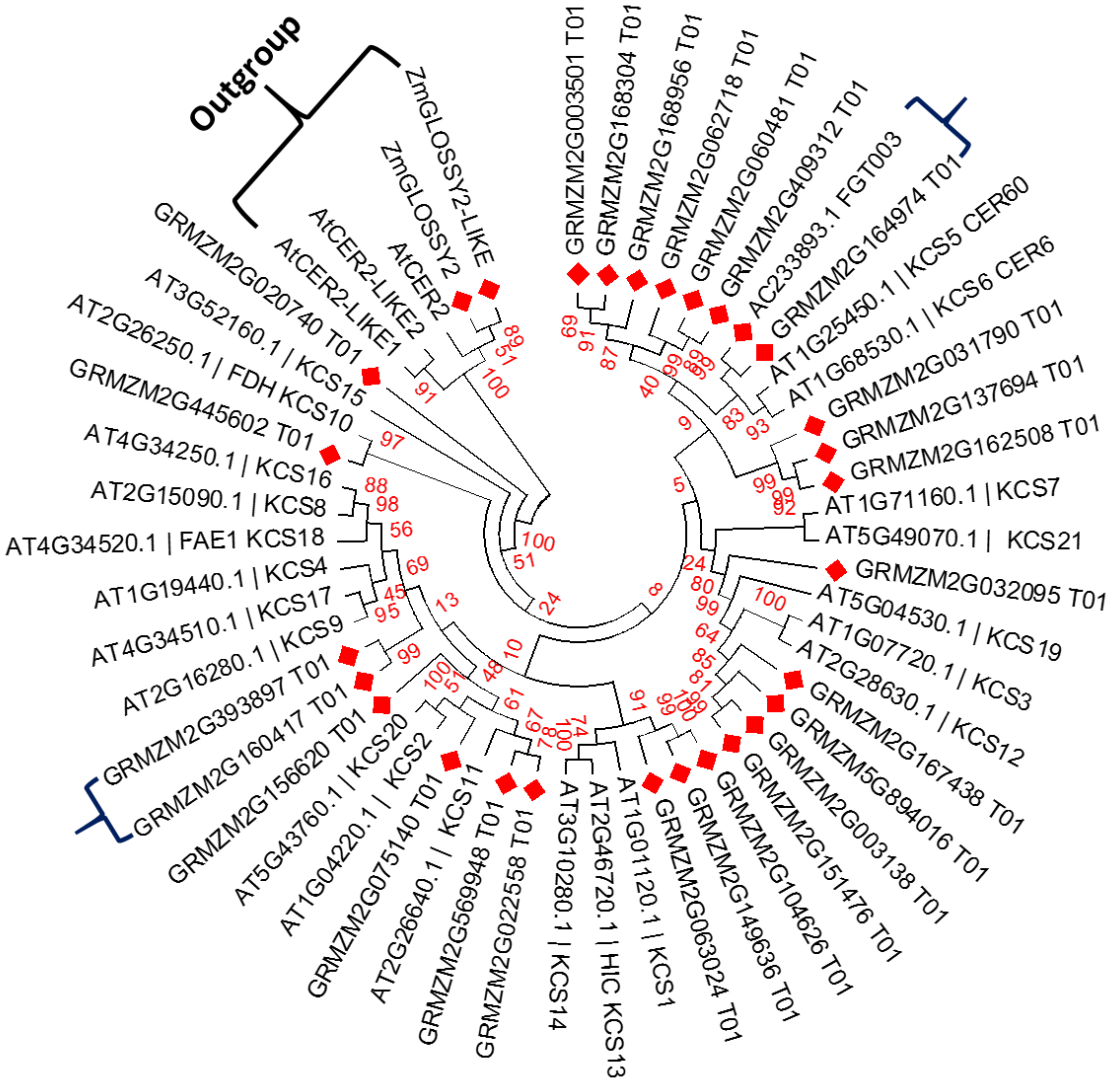


Figure 1. Phylogenetic relationship among maize KCS and Arabidopsis KCS based on their amino acid sequence. MEGA v 6.0 was used for the construction of the phylogenetic neighbor-joining tree. Red diamonds denote the maize KCS and the blue brackets denote the maize KCS used in this study

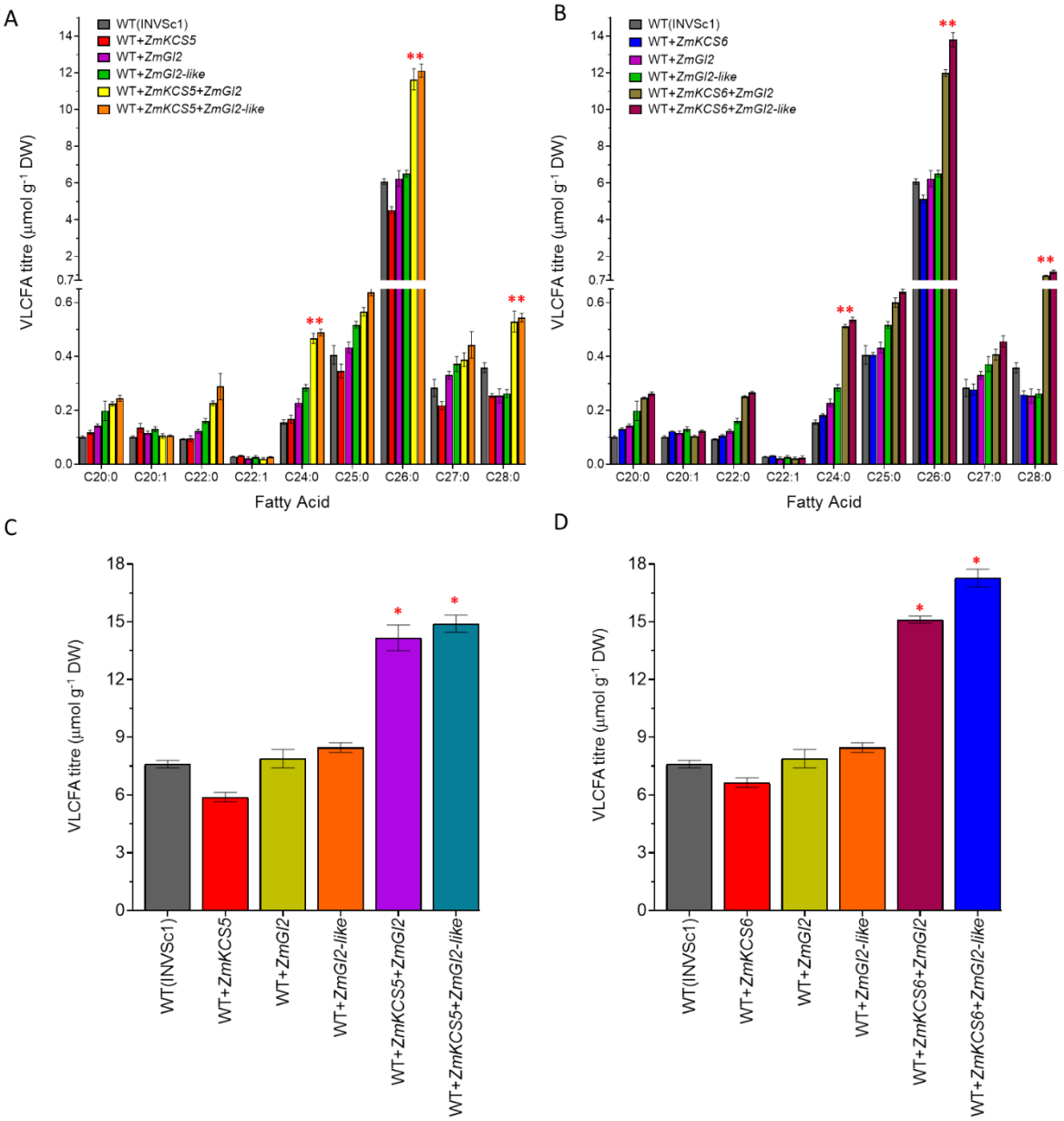


Figure 2. VLCFA composition (A,B) and total VLCFA accumulation (C,D) of yeast INVSc1 strains co-expressing either maize KCS5 (A,C) or KCS6 (B,D) and *Glossy2* gene family under GAL inducible promoter, n=6 replicates per genotype. Statistical significance as compared to WT yeast was determined using a student's t-test, where * ≤ p-value of 0.05.

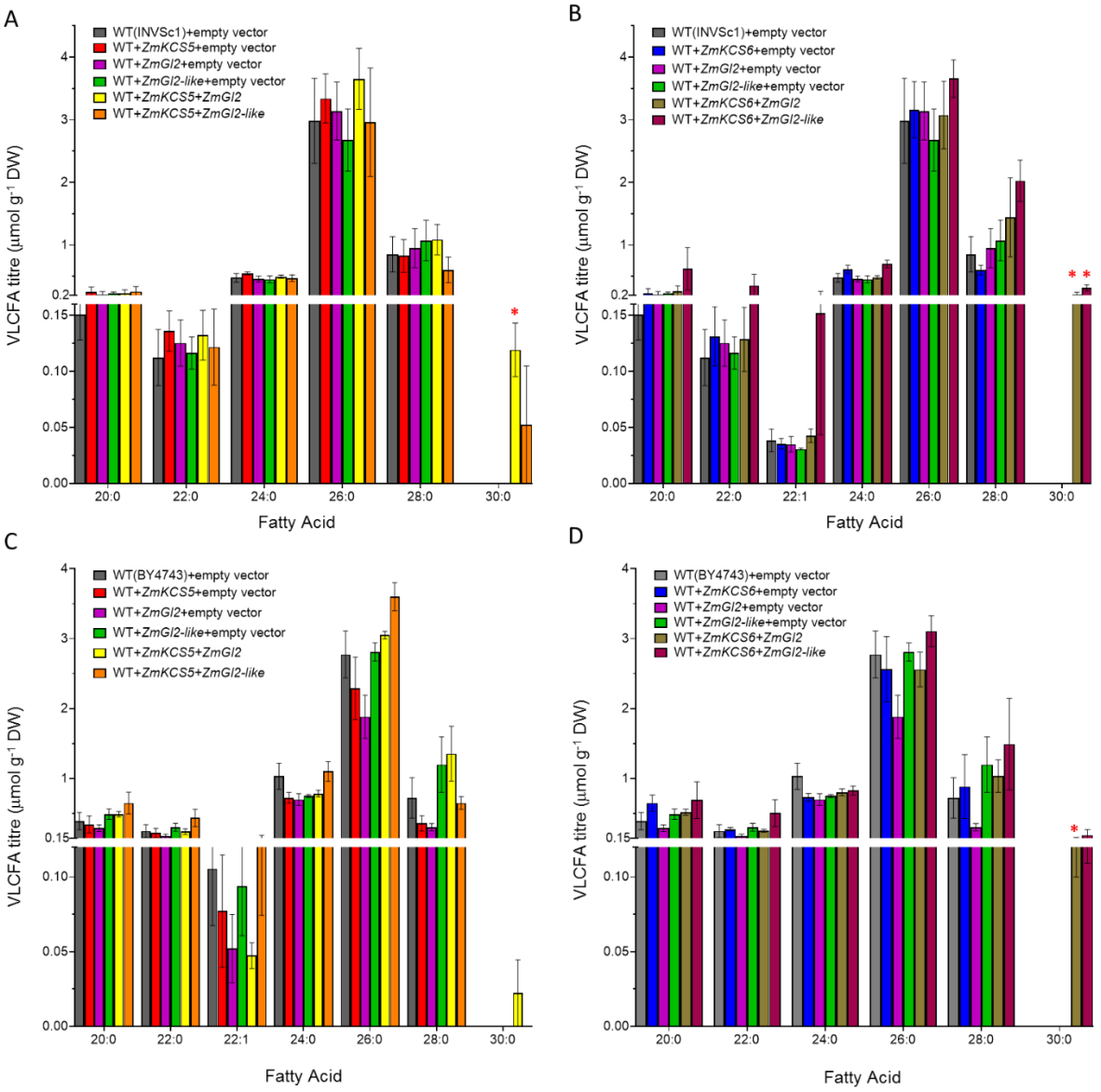


Figure 3. VLCFA composition of yeast INVSc1 (A, B) and BY4743 (C, D) strains co-expressing either maize KCS 5(A,C) or KCS6 (B,D) and *Glossy2* gene family under GPD inducible promoter, n=3 replicates per genotype. Statistical significance as compared to WT yeast was determined using a student's t-test, where * ≤ p-value of 0.05.

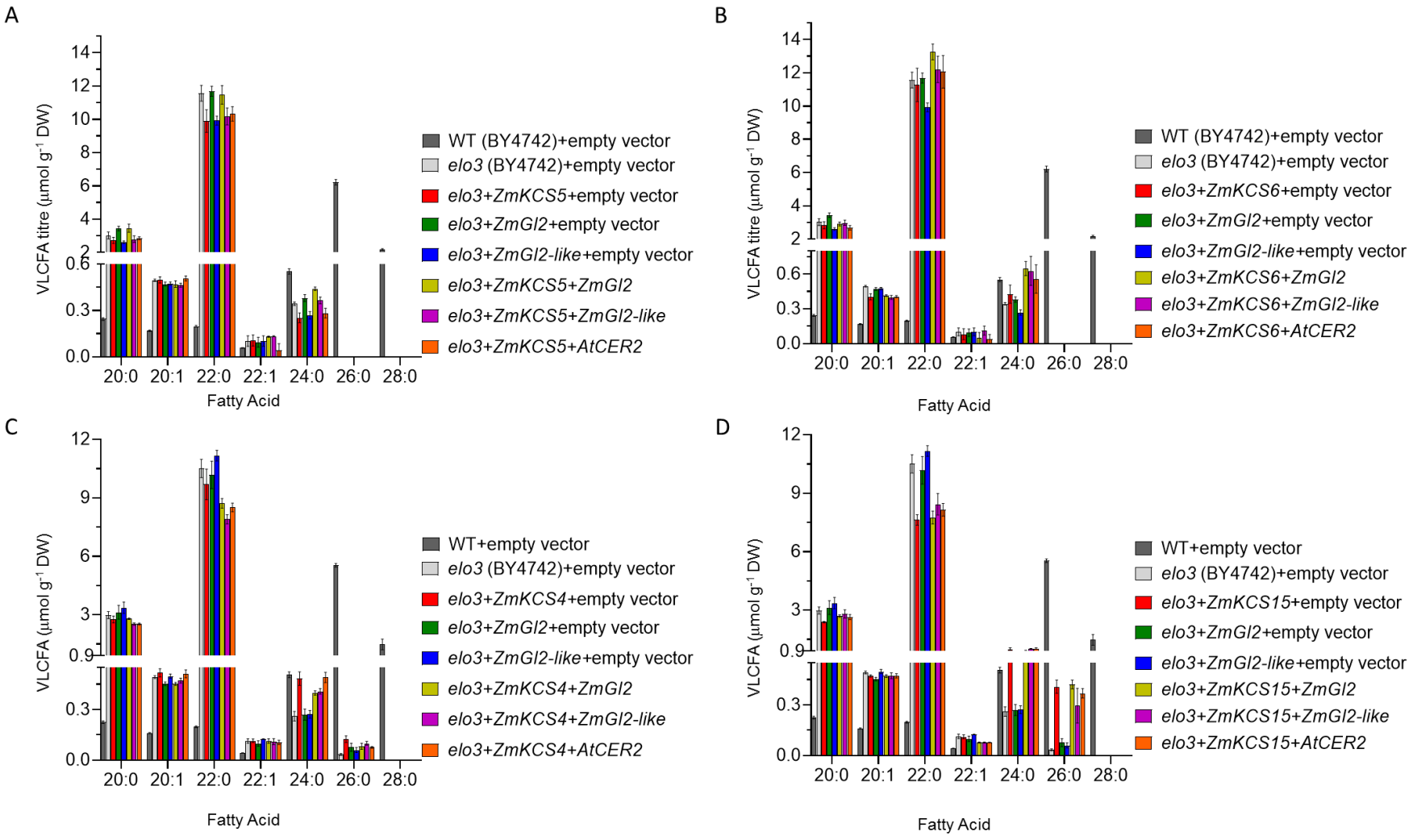


Figure 4. VLCFA composition of yeast *elo3* mutant strains co-expressing either the maize KCS5 (A), KCS6 (B), KCS4 (C), or KCS15 (D) and with either *ZmGl2*, *ZmGl2-like* or *AtCER2* genes, all under the control of the GPD constitutive promoter, Data are the average ± standard error (n=4 replicates per genotype).

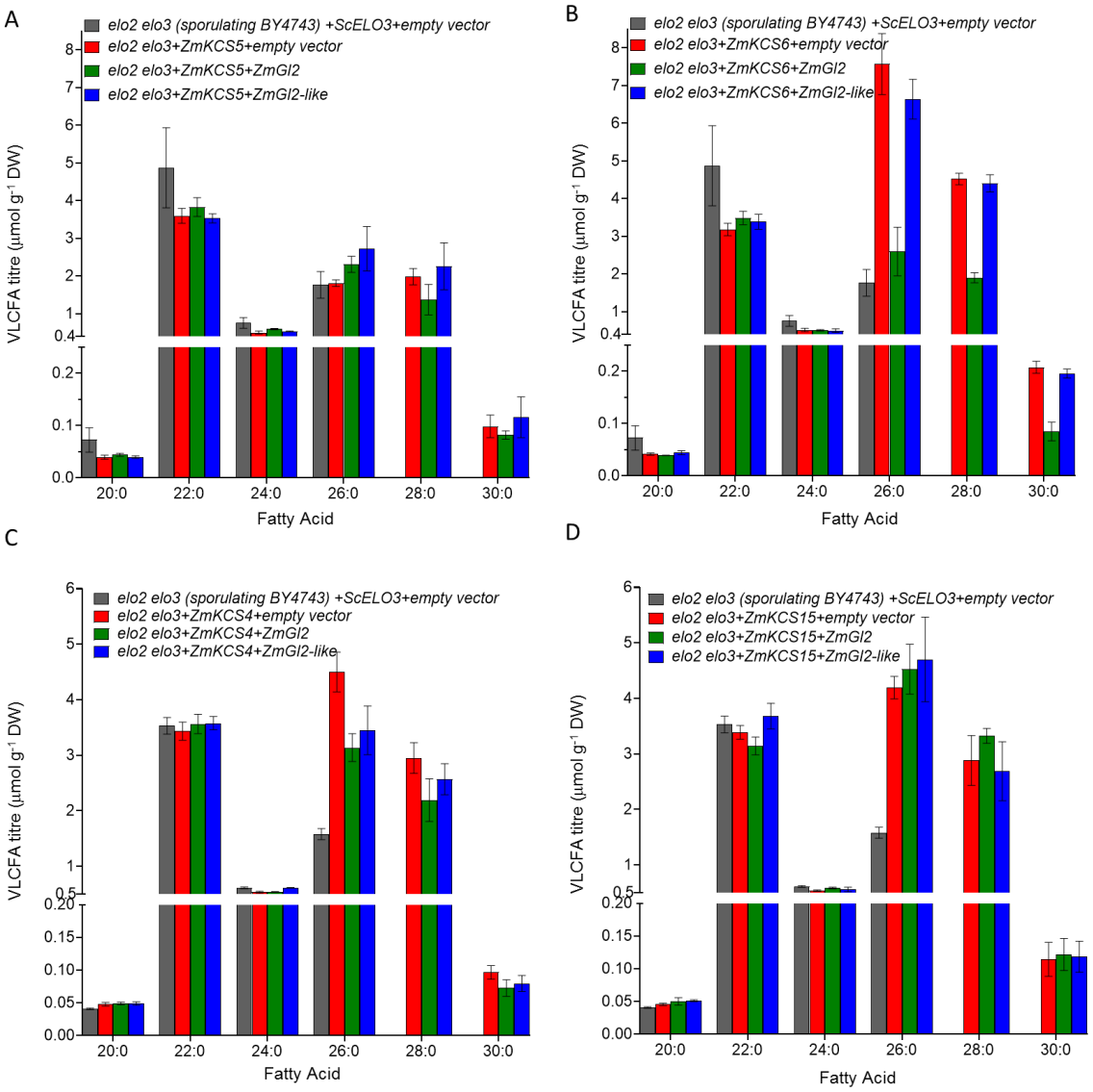


Figure 5. VLCFA composition of yeast *elo2elo3* double mutant strains co-expressing either maize KCS5 (A), KCS6 (B), KCS4 (C) or KCS15 (D) with *Glossy2* gene family. All genes are under the transcriptional control of the GPD constitutive-promoter. Data are the average ± standard error (n=3 replicates per genotype).

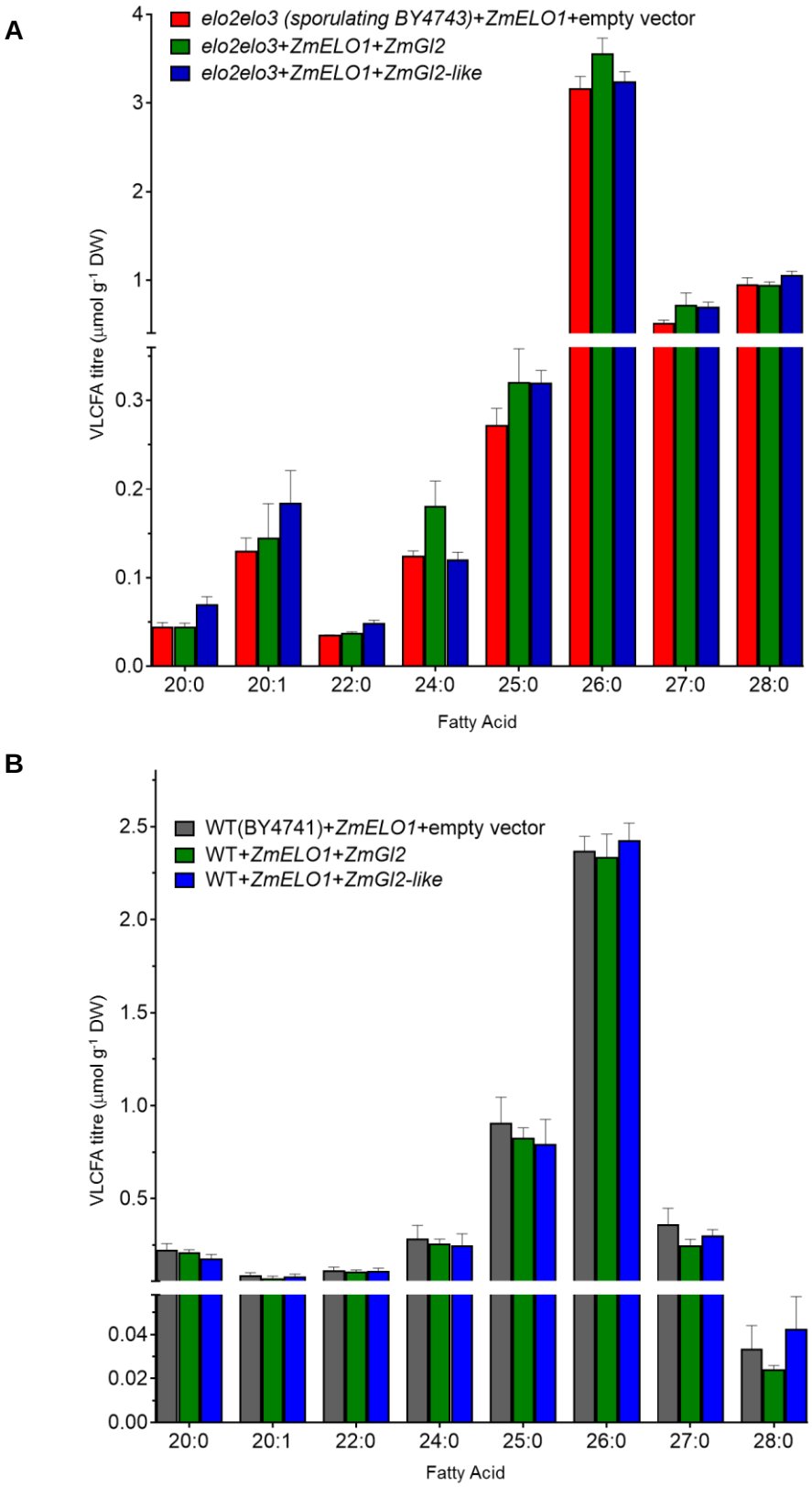


Figure 6. VLCFA composition of yeast *elo2elo3* double mutant strains co-expressing the maize ELO1 with either the ZmGl2 or ZmGl2-like. All genes are under the transcriptional control of the *GAL1* inducible-promoter. Data are the average ± standard error (n=4 replicates per genotype).

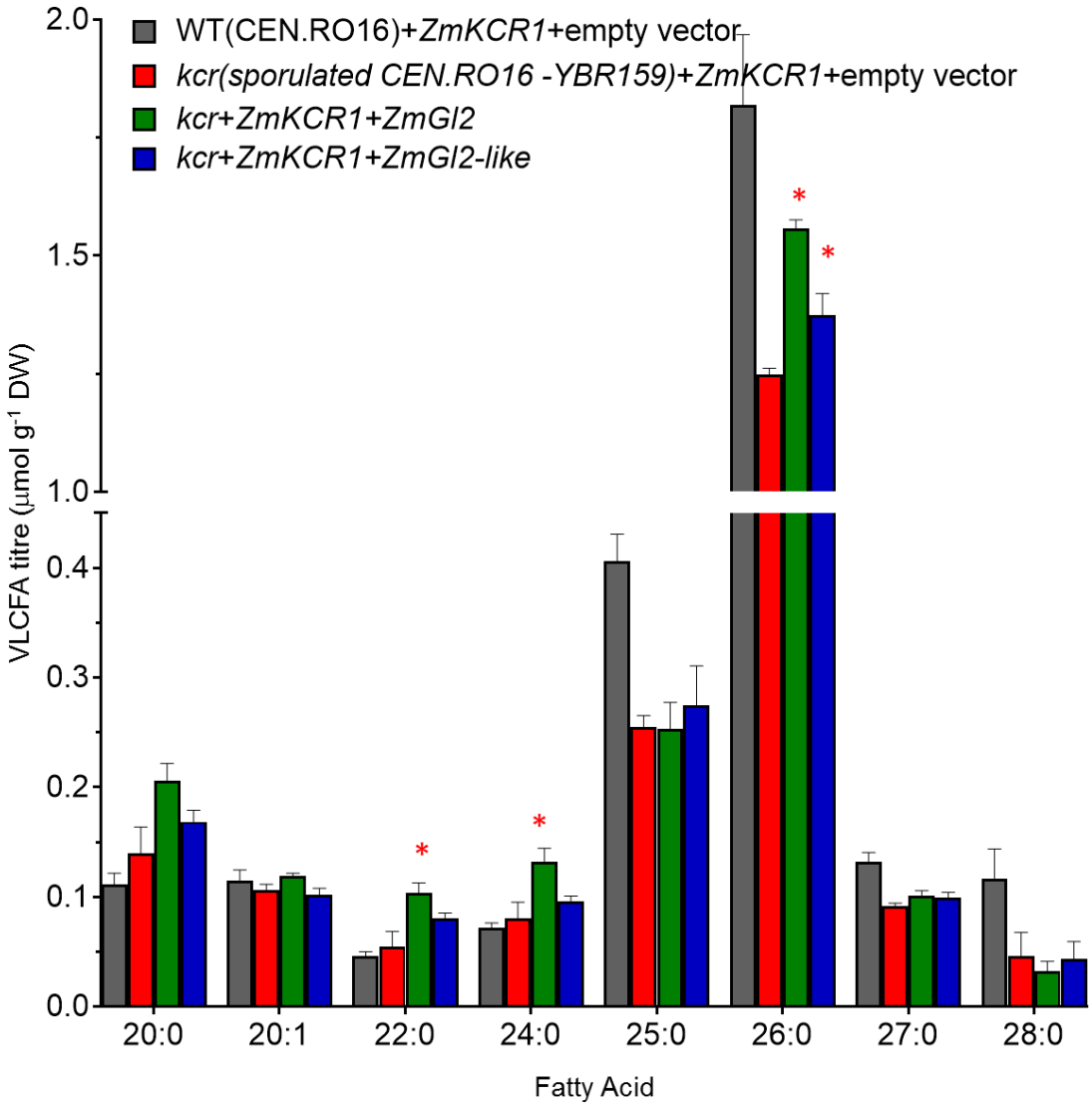


Figure 7. VLCFA composition of yeast *kcr* mutant strain, co-expressing the maize *KCR1* gene with either the ZmGl2 or ZmGl2-like genes. All genes are under the transcriptional control of the *GAL1* inducible-promoter. Data are the average ± standard error (n=4 replicates per genotype); asterisks indicate samples that are statistically different from the *kcr* mutant strain using a student's t-test; p-value ≤ 0.05.

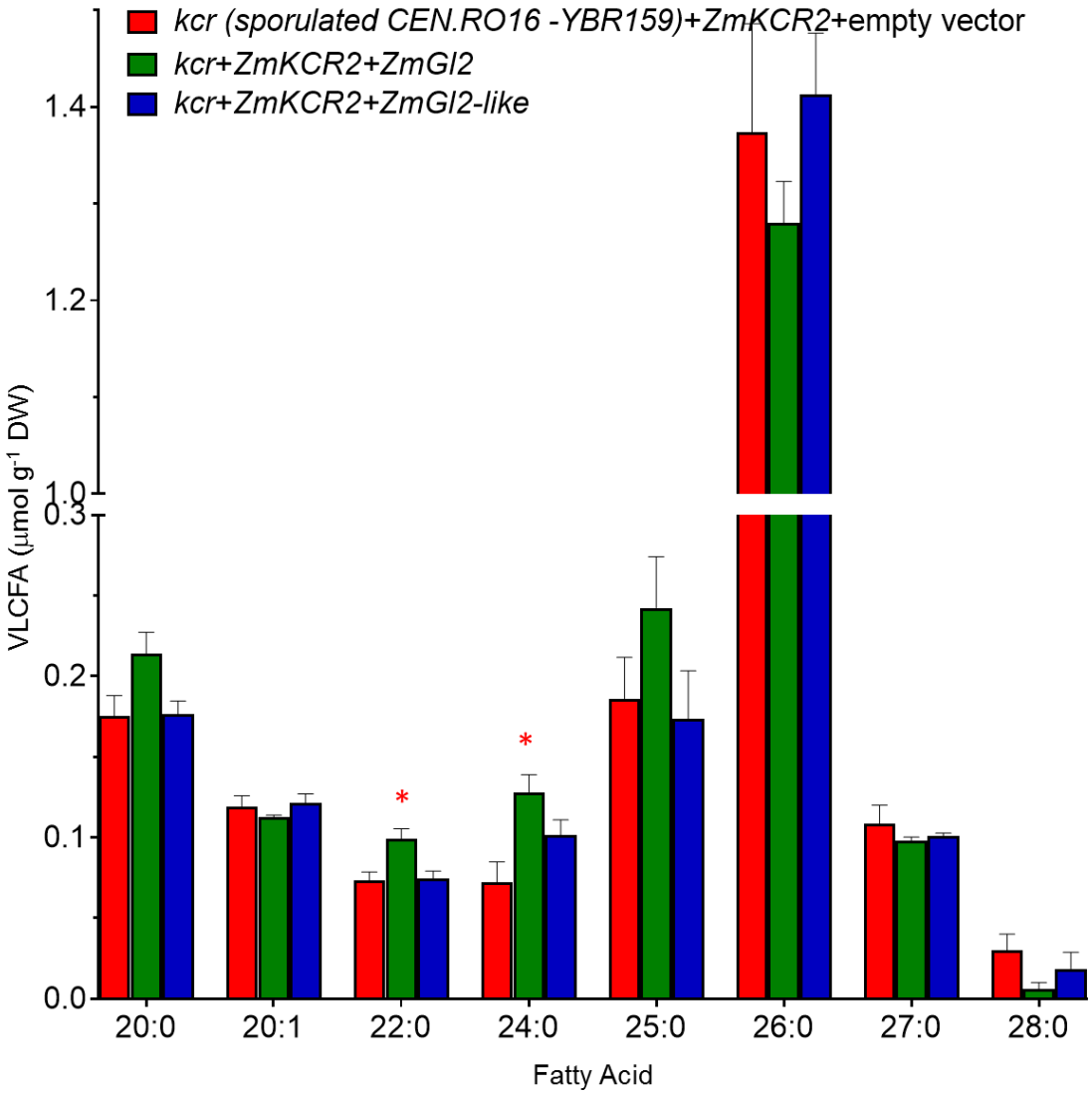


Figure 8. VLCFA composition of yeast *kcr* mutant strain, co-expressing the maize *KCR2* gene with either the ZmGl2 or ZmGl2-like genes. All genes are under the transcriptional control of the *GAL1* inducible-promoter. Data are the average ± standard error (n=4 replicates per genotype); asterisks indicate samples that are statistically different from the *kcr* mutant strain using a student's t-test; p-value ≤ 0.05.

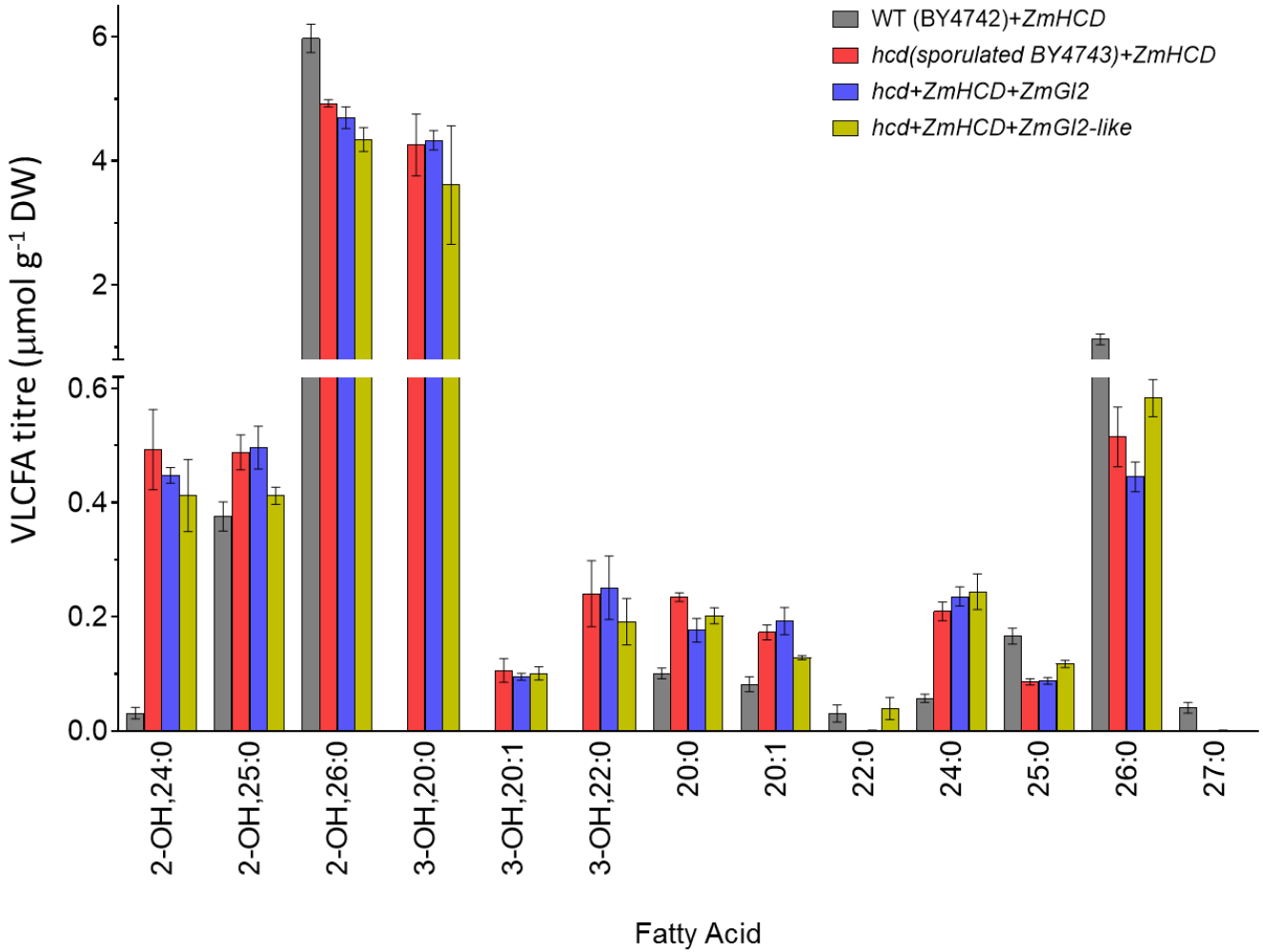
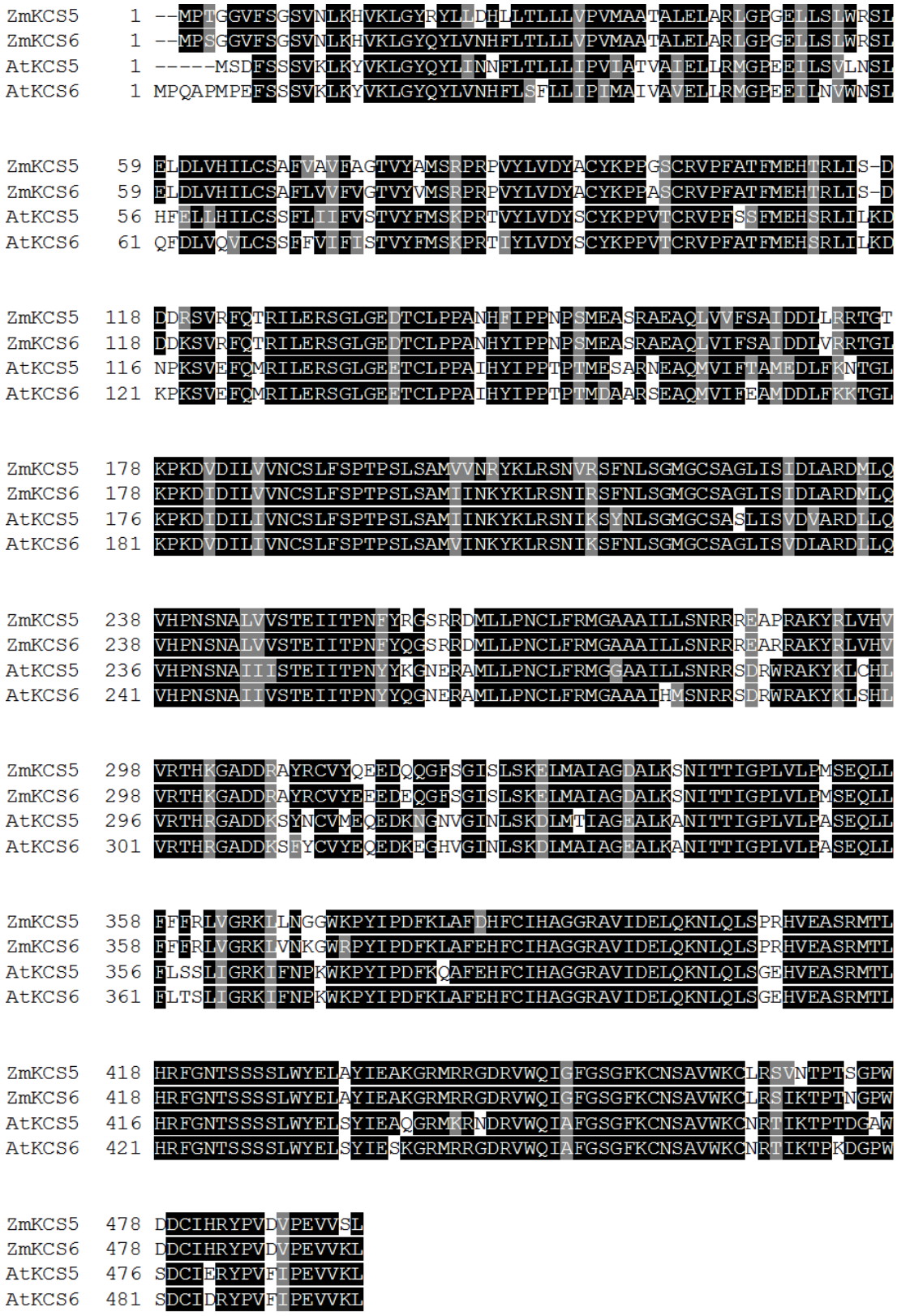


Figure 9. VLCFA composition of individual VLCFA products of the yeast *hcd* mutant strain, expressing ZmHCD, co-expressed with either the *Glossy2* or *Glossy2-like* gene. All genes are under the transcriptional control of the *GAL1* inducible-promoter. Data are the average ± standard error (n=4 replicates per genotype)



Supplemental Figure 1. Amino acid sequence alignment of *Zea mays* CER6 homologs. The maize KCS5 (AC233893.1_FGT003) and KCS6(GRMZM2G164974) sequences were aligned with the Arabidopsis CER6/KCS6 (At1g68530), CER60/KCS5 (At1g25450) sequences using Clustal O (1.2.4) and BOXSHADE (v3.21).Black shading identifies identical residues, and gray-shading identifies similar residues.

CHAPTER 8. GENERAL CONCLUSION

This dissertation discusses multiple applications of mass spectrometry technologies within the field of plant metabolomics, focusing on mass spectrometry imaging tools - MALDI and ELDI. These technologies have been used to investigate metabolite trafficking which is key to understanding metabolic networks. Four biological applications were investigated, 1) the metabolic process of maize seed germination as influenced by genetic differences in inbred lines - B73 and Mo17, 2) the distribution of membrane lipids on the photosynthetic machinery of C4-plants as influenced by genetic disposition of maize inbreds B73 and Mo17 and their reciprocal hybrids, 3) the effect of external stimuli, light, on anthocyanin-rich coleus plants, and 4) the distribution of plant epicuticular surface lipids on Arabidopsis flowers that express maize transgenes *Glossy2* and *Glossy2-like*. Traditional metabolic strategies employ the use of ground tissues thereby losing spatial information, these mass spectral imaging technologies combined with LC- and GC- mass spectrometry have provided a holistic outlook on metabolic networks as they occur within single cells and tissues. Also described is the characterization of maize genes *Glossy2* and *Glossy2-like* in plant epicuticular lipid biogenesis.

Germination is a highly regulated compartmentalized process

The process of germination is dependent on the taxonomic clade of the plant, where monocots such as maize have a starch- and protein-filled endosperm and a lipid-filled embryo, surrounded by an aleurone layer that houses the enzymes essential for the process of germination (Dante et al., 2015). Experimental results in Chapter 2 demonstrated that the metabolic network during germination is quantitatively and spatially influenced by genotypic and developmental differences as determined by a combination of MALDI-MSI and LC- and GC- MS. Interestingly, the genetic diversity influenced the distribution free aminoacids (Thr, Lys, Arg and Asn) and the

accumulation of organic acids (malate, pyruvate and fumarate), where the former increased in Mo17 and the latter in B73 inbreds. The high spatial capabilities, chemical versatility, and increased sensitivity of the MALDI-MSI technology demonstrated that during early germination ceramides distribute to the endosperm-scutellar boundary, large polysaccharides localize to the endosperm, and fatty acids and lipid species (phospholipids and TAGs) localize in the embryo. Within the embryo, phospholipids and TAGs showed a preferential distribution of fatty acid species that depended upon the degree of unsaturation, where oleate (18:1) was absent in the radicle whereas, linoleate (18:2) homogenously distributed in the embryo. The enrichment of 18:2-containing lipids in the radicle correlated with the differential expression patterns of the FAD2 gene which was increased in the radicle tissue relative to scutellar tissue as determined by qPCR. The combination of these strategies thereby provided new insights into the germination processes which integrates reactions across cellular compartments.

Thylakoid membrane lipids are developmentally and genetically controlled

Although physical separation of different cell types is possible through lipidomic, transcriptomic or proteomic analysis, these methods are prone to undesirable degradation or cross-contamination of the materials during the sample preparation process. Chapter 3 demonstrated the high spatial resolution capabilities of MALDI-MSI (5μm) in visualizing metabolite distributions within intact tissues and cellular environments. This was demonstrated on maize leaves that are unique for their Kranz anatomy, where photoautotrophic photosynthesis is differentiated between two photosynthetic cells, the bundle sheath (BS) and the mesophyll (M) (Edwards *et al.* 2001, Hatch 1987, von Caemmerer and Furbank 2003). Both these cells differ in their chloroplast-metabolic capabilities and ultrastructures, where in contrast to M cells, BS cells contain Rubisco and fix CO_2, and lack photosystem II (PS II) and granal stacks (Woo *et al.* 1970). MALDI-MSI

demonstrated that the thylakoid membrane lipids, sulfoquinovosyldiacylglycerols (SQDG) and phosphatidylglycerols (PG), differed in BS and M cells. MALDI-MSI also demonstrated that the SQDG and PG distribution depended on the genetic composition of the plant. SQDG species showed uniform distribution in both M and BS cells regardless of genotype and leaf development, however the PG distribution depended on the genotype and the fatty acyl chain constituent, where M cells are predominantly 16:1-containing PGs and BS chloroplasts mostly 16:0-containing PGs. Genotype-specific differences in cellular distribution was evident for PG 32:0 with a preference in localization for B73 BS cells and a more uniform distribution between BS and M cells in Mo17. Interestingly, the reciprocal hybrids exhibited maternal inheritance such that the localization of PG 32:0 in B73xMo17 resembled the distribution in the B73 parental inbred and Mo17xB73 resembled the Mo17 parent. This study provided unprecedented insights to metabolic outcomes at single-cell resolution.

Plant UV-absorbing molecules provide a pseudo-matrix for matrix-free mass spectral imaging technologies

The above two applications are conducted under reduced pressure ionization conditions and requires the application of matrices. Chapter 4 demonstrated the application of mass spectral imaging in a matrix-free ambient air ionization environment via ELDI. This technology was applied to determine the spatial distribution of metabolites in Coleus Henna leaves that were grown with only half of the leaf illuminated. These leaves show a chartreuse green to copper color that changes with surrounding conditions such as light, and temperature (Lebowitz, 2011). Since plants accumulate large quantities of UV-absorbing molecules the plant itself served as a pseudo-matrix that enabled ELDI to visualize both UV-absorbing metabolites (flavonoids, terpenes, phenolics) and non-UV absorbers (sugars, amino acids and organic acids). Experimental results in Chapter 4

revealed that major metabolic changes were observed for anthocyanins, where cyanidin glycosides were the most abundant on the illuminated side. Additionally, this technology was powerful in detecting the spatial distribution of structurally related metabolites, cyanidin-coumaroylglucoside-glucoside, cyanidin-malonyl glucoside, and cyanidin. Amongst the non-UV absorbers significant differences were observed for sugars. The experimental design of shade treatment resulted in the detection of novel source-sink relationships observed between the 2 halves of the leaves, where the fully-illuminated half served as the source tissue that fixed carbon primarily in the form of sugars compensating for the reduced photosynthesis imposed by shading the other half of the leaf. ELDI-MSI is in the early stages of application to plant biology, our study demonstrated that this technology is capable of spatially identifying flavonoids, carbohydrates, aminoacids, and organic acids, but detection of lipids was limited.

Maize Glossy2 and Glossy2-like genes impact extracellular epicuticular lipid accumulation

Chapters 5, 6, and 7 explored the role of maize *Glossy2* (*Gl2*) and *Glossy2-like* (*Gl2-like*) genes in epicuticular surface lipid biogenesis. Using transgenic approaches, these studies have provided a new role to an uncharacterized *Gl2-like* gene, demonstrating its role in epicuticular biosynthesis. Using extensive metabolomics approaches we have demonstrated that *Gl2* and *Gl2-like* are functional homologs of Arabidopsis CER2, however differing in their lipid constituents, and *Gl2, Gl2-like* and *CER2* are essential for the purpose of elongation beginning at C26 chain length. The majority of the past genetic based studies of the accumulation of epicuticular lipids are primarily limited to a single time-point of analysis, and overlook the dynamics of the processes that integrate the developmental program of the tissue/organs that are being evaluated (Jenks et al., 1995; Goodwin et al., 2005; Pascal et al., 2013). Furthermore, most such data are collected from the analysis of epicuticular lipid extracts, which are prepared from relatively large biological

samples that combine lipids generated by a large collection of cells that are most probably of different age and different developmental states. Emerging flowers and the application of analytical mass spectrometry technologies; PTV-GC-MS and MALDI-MSI, offered the potential to address both of these potential limitations in the deciphering epicuticular lipid accumulation patterns. The comparisons between PTV-GC-MS and MALDI-MSI revealed that the process of epicuticular lipid biogenesis is dynamic and *Gl2, Gl2-like* and *CER2* impacted the distribution of epicuticular lipids on the surface of developing flowers. MALDI-MSI demonstrated that epicuticular lipids are more abundant on the abaxial floral surfaces, and the process of fatty acid elongation localizes predominantly in the petals of the flower, whereas reductive and decarboxylative steps of epicuticular biogenesis occurs predominantly in the sepals.

Because epicuticular lipids biosynthesis occurs within epidermal cells, an extensive analysis of the lipidome was analyzed using LC-qTOF. Lipidomic analysis demonstrated *Gl2, Gl2-like* and *CER2* altered the levels of very long chain fatty acids starting at acyl chains 26:0. Since fatty acid elongation process is carried out by the FAE complex, the mechanistic understanding of the interactions between *Gl2* and *Gl2-like* with either KCS or FAE components, was addressed using co-expression in yeast heterologous systems. These studies demonstrated that *Gl2* and *Gl2-like* interacted with 2 out of 4 KCS tested, by the detection of 30:0 fatty acid which was absent when the maize genes are expressed individually. However, this finding was dependent on the yeast strain used and could not be recapitulated in yeast strains that are devoid of native yeast FAE components. Such strains are lethal but can be maintained by genetic complementation of maize KCS (Campbell et al., 2019). Additionally, the yeast expression studies also demonstrated that co-expression of *Gl2* and *Gl2-like* with maize KCR1 and KCR2 altered the fatty acid constituents of yeast however without the formation of new products. Based on amino acid sequence similarities,

GL2, GL2-LIKE and CER2 proteins are classified as being members of subclade-II BAHD acyltransferase family (D'auria, 2006). Using site-directed mutagenesis strategies, we demonstrated that GL2 and GL2-LIKE mutated at their BAHD- catalytic motif differed in their function, where only GL2-LIKE required the BAHD catalytic motif for epicuticular lipid biogenesis.

To conclude, this research has generated new insights to gene functionality which could have implications on the understanding of epicuticular lipid biosynthesis. Nevertheless the mechanism by which both GL2 and GL2-LIKE proteins modulate the fatty acid elongation process, needs further investigation. The stoichiometry of how the FAE complex exist is unknown, this make deciphering the interactions complicated. It is possible to speculate that these genes may regulate specific iterations of the FAE elongation cycles, however answering these questions would require development of direct biochemical assays that go beyond the studies to date, which rely on correlations between genetic modifications and metabolic outcomes. We have demonstrated that GL2 is a soluble protein when expressed in *E.coli* (Appendix B), this can be used to explore structure and function of the protein, which can provide insights concerning its *in vivo* function.

References

Amstalden van Hove, E.R., Smith, D.F., Fornai, L., Glunde, K., Heeren, R.M.A., 2011. An Alternative Paper Based Tissue Washing Method for Mass Spectrometry Imaging: Localized Washing and Fragile Tissue Analysis. J. Am. Soc. Mass Spectrom. 22, 1885–1890. https://doi.org/10.1007/s13361-011-0203-z

Angel, P.M., Caprioli, R.M., 2013. Matrix-Assisted Laser Desorption Ionization Imaging Mass Spectrometry: In Situ Molecular Mapping. Biochemistry 52, 10.1021/bi301519p. https://doi.org/10.1021/bi301519p

Benning, C., 2008. A role for lipid trafficking in chloroplast biogenesis. Prog. Lipid Res. 47, 381–389. https://doi.org/10.1016/j.plipres.2008.04.001

Bernard, A., Joubès, J., 2013. Arabidopsis cuticular waxes: Advances in synthesis, export and regulation. Prog. Lipid Res. 52, 110–129. https://doi.org/10.1016/j.plipres.2012.10.002

Bianchi, 1975. Glossy mutants of maize. VI. Chemical constituents of glossy-2 epicuticular waxes. Maydica 20, 165–173.

Bianchi Rome (Italy)), A. (Istituto S. per la C., Bianchi, G., Avato, P. (Pavia U. (Italy). D. di C.O., Salamini Koeln (Germany, F.R.). Erwin-Baur-Inst.), F. (Max-P.I. fuer Z., 1985. Biosynthetic pathways of epicuticular wax of maize as assessed by mutation, light, plant age and inhibitor studies. Maydica (Italy).

Boughton, B.A., Thinagaran, D., Sarabia, D., Bacic, A., Roessner, U., 2016. Mass spectrometry imaging for plant biology: a review. Phytochem. Rev. 15, 445–488. https://doi.org/10.1007/s11101-015-9440-2

Campbell AA, Stenback KE, Flyckt K, Hoang T, Perera MADN, Nikolau BJ. A single-cell platform for reconstituting and characterizing fatty acid elongase component enzymes. Submitted to PLoSOne 2019

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., Prasher, D.C., 1994. Green Fluorescent Protein as a Marker for Gene Expression. Science. 263, 802–805. https://doi.org/10.1126/science.8303295

D'Auria, J.C., 2006. Acyltransferases in plants: a good time to be BAHD. Curr. Opin. Plant Biol. 9, 331–340. https://doi.org/10.1016/j.pbi.2006.03.016

Dante RA, Larkins BA, Sabelli PA (2015) Cell cycle control and seed development. 832 Advances in Seed Biology: 21

DeBono A, et al. (2009) Arabidopsis LTPG Is a Glycosylphosphatidylinositol-Anchored Lipid Transfer Protein Required for Export of Lipids to the Plant Surface. The Plant Cell 21(4):1230-1238.

Edwards, G.E., Franceschi, V.R., Ku, M.S.B., Voznesenskaya, E. V, Pyankov, V.I., Andreo, C.S., 2001. Compartmentation of photosynthesis in cells and tissues of C 4 plants. J. Exp. Bot. 52, 577–590. https://doi.org/10.1093/jexbot/52.356.577

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grun, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R.B., Briggs, S.P., 1997. Analysis of a chemical plant defense mechanism in grasses. Anal. a Chem. plant Def. Mech. grasses 696–699.

Gallagher, S.R. (1992) GUS Protocols: Using the GUS Gene as A Reporter of Gene Expression. New York: Academic Press.

Glauser, G., Marti, G., Villard, N., Doyen, G.A., Wolfender, J.-L., Turlings, T.C.J., Erb, M., 2011. Induction and detoxification of maize 1,4-benzoxazin-3-ones by insect herbivores.(Report). Plant J. 68, 901. https://doi.org/10.1111/j.1365-313X.2011.04740.x

Goodwin, S.M., Rashotte, A.M., Rahman, M., Feldmann, K.A., Jenks, M.A., 2005. Wax constituents on the inflorescence stems of double eceriferum mutants in Arabidopsis reveal complex gene interactions. Phytochemistry 66, 771–780. https://doi.org/10.1016/j.phytochem.2005.02.001

Griffin, B.A., Adams, S.R., Tsien, R.Y., 1998. Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. Science (80-.). 281, 269–272. https://doi.org/10.1126/science.281.5374.269

Haslam, T.M., Mañas-Fernández, A., Zhao, L., Kunst, L., 2012. Arabidopsis ECERIFERUM2 is a component of the fatty acid elongation machinery required for fatty acid extension to exceptional lengths. Plant Physiol. 160, 1164–74. https://doi.org/10.1104/pp.112.201640

Hatch, M.D., 1987. C 4 photosynthesis: a unique elend of modified biochemistry, anatomy and ultrastructure. BBA Rev. Bioenerg. 895, 81–106. https://doi.org/10.1016/S0304-4173(87)80009-5

Hayes, H.K., Brewbaker, H.E., 1928. Glossy Seedlings in Maize. Am. Nat. 62, 228–235.

Huang, M.-Z., Cheng, S.-C., Jhang, S.-S., Chou, C.-C., Cheng, C.-N., Shiea, J., Popov, I.A., Nikolaev, E.N., 2012. Ambient molecular imaging of dry fungus surface by electrospray laser desorption ionization mass spectrometry. Int. J. Mass Spectrom. 325–327, 172–182. https://doi.org/https://doi.org/10.1016/j.ijms.2012.06.015

Jellings, A.J., Leech, R.M., 1982. The Importance of Quantitative Anatomy in the Interpretation of Whole Leaf Biochemistry in Species of Triticum, hordeum and Avena. New Phytol. 92, 39–48.

Jenks, M. A., Tuttle, H. A., Eigenbrode, S.D., Feldmann, K. A., 1995. Leaf Epicuticular Waxes of the Eceriferum Mutants in Arabidopsis. Plant Physiol. 108, 369–377. https://doi.org/10.1104/pp.108.1.369

Jenks, M.A., Joly, R.J., Peters, P.J., Rich, P.J., Axtell, J.D., Ashworth, E.N., 1994. Chemically Induced Cuticle Mutation Affecting Epidermal Conductance to Water Vapor and Disease Susceptibility in Sorghum bicolor (L.) Moench. Plant Physiol. 105, 1239–1245. https://doi.org/10.1104/pp.105.4.1239

Jentaie, S., Min-Zon, H., Hsiu-Jung, Hs., Chi-Yang, L., Cheng-Hui, Y., Iwona, B., Jan, S., 2005. Electrospray-assisted laser desorption/ionization mass spectrometry for direct ambient analysis of solids. Rapid Commun. Mass Spectrom. 19, 3701–3704. https://doi.org/10.1002/rcm.2243

Jetter, R., Schäffer, S., Riederer, M., 2000. Leaf cuticular waxes are arranged in chemically and mechanically distinct layers: evidence from Prunus laurocerasus L. Plant. Cell Environ. 23, 619–628. https://doi.org/10.1046/j.1365-3040.2000.00581.x

Jungmann, J.H., Heeren, R.M.A., 2012. Emerging technologies in mass spectrometry imaging. J. Proteomics 75, 5077–5092. https://doi.org/10.1016/j.jprot.2012.03.022

Kerstiens, G., 1996. Cuticular water permeability and its physiological significance. J. Exp. Bot. 47, 1813–1832.

Klein, A.T., Yagnik, G.B., Hohenstein, J.D., Ji, Z., Zi, J., Reichert, M.D., MacIntosh, G.C., Yang, B., Peters, R.J., Vela, J., Lee, Y.J., 2015. Investigation of the Chemical Interface in the Soybean–Aphid and Rice–Bacteria Interactions Using MALDI-Mass Spectrometry Imaging. Anal. Chem. 87, 5294–5301. https://doi.org/10.1021/acs.analchem.5b00459

Kompauer, M., Heiles, S., Spengler, B., 2017. Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4-mum lateral resolution. Nat. Methods 14, 90–96. https://doi.org/10.1038/nmeth.4071

Kolattukudy, P.E., 1965. Biosynthesis of Wax in Brassica oleracea *. Biochemistry 4, 1844–1855. https://doi.org/10.1021/bi00885a023

Kolattukudy, P.E., 1985. Enzymatic Penetration of the Plant Cuticle by Fungal Pathogens. Annu. Rev. Phytopathol. 23, 223–250. https://doi.org/10.1146/annurev.py.23.090185.001255

Kolattukudy, P.E., Croteau, R., and Buckner, J.S., 1976. Biochemistry of plant waxes, in P.E. Kolattukudy (ed.), Chemistry and Biochemistry of Natural Waxes, Elsevier Press, New York, pp. 289 - 347.

Koo, J.., Kim, Y., Kim, J., Yeom, M., Lee, I.C., Nam, H.G., 2007. A GUS/luciferase fusion reporter for plant gene trapping and for assay of promoter activity with luciferin-dependent control of the reporter protein stability. A GUS/luciferase fusion Report. plant gene Trapp. assay Promot. Act. with luciferin-dependent Control Report. protein Stab. 1121–1131.

Koornneef, M., Hanhart, C.J., Thiel, F., 1989. A Genetic and Phenotypic Description of Eceriferum (cer) Mutants in Arabidopsis thaliana. J. Hered. 80, 118–122.

Korte, A.R., Yandeau-Nelson, M.D., Nikolau, B.J., Lee, Y.-J., 2015. Subcellular-level resolution MALDI-MS imaging of maize leaf metabolites by MALDI-linear ion trap-Orbitrap mass spectrometer.

Küpper, H., Seib, L.O., Sivaguru, M., Hoekenga, O.A., Kochian, L. V, 2007. A method for cellular localization of gene expression via quantitative in situ hybridization in plants: Quantitative in situ hybridization. Plant J. 50, 159–187. https://doi.org/10.1111/j.1365-313X.2007.03031.x

Lebowitz, R. J. (2011, February 9). The Genetics and Breeding of Coleus. Plant Breeding 604 Reviews. http://doi.org/doi:10.1002/9781118061008.ch9

Lee, Y.J., Perdian, D.C., Song, Z., Yeung, E.S., Nikolau, B.J., Jin, L.Y., C., P.D., Zhihong, S., S., Y.E., J., N.B., 2012. Use of mass spectrometry for imaging metabolites in plants. Plant J. 70, 81–95. https://doi.org/10.1111/j.1365-313X.2012.04899.x

Leide, J., Hildebrandt, U., Reussing, K., Riederer, M., Vogg, G., 2007. The Developmental Pattern of Tomato Fruit Wax Accumulation and Its Impact on Cuticular Transpiration Barrier Properties: Effects of a Deficiency in a β-Ketoacyl-Coenzyme A Synthase (LeCER6). Plant Physiol. 144, 1667 LP-1679. https://doi.org/10.1104/pp.107.099481

Long, L.M., Patel, H.P., Cory, W.C., Stapleton, A.E., 2003. The maize epicuticular wax layer provides UV protection. Funct. Plant Biol. 30, 75–81.

Loneman, D.M., Peddicord, L., Al-Rashid, A., Nikolau, B.J., Lauter, N., Yandeau-Nelson, M.D., 2017. A robust and efficient method for the extraction of plant extracellular surface lipids as applied to the analysis of silks and seedling leaves of maize. PLoS One 12, e0180850.

Martin JT, J.B. 1970., 1970. The Cuticles of plants: J.T. Martin and B.E. Juniper. Edward Arnold (Publishers) Ltd., London and New York, pp. x+347, 1970. (Book Review). Micron (1969). https://doi.org/10.1016/0047-7206(70)90020-8

McDonnell, L.A., Heeren, R.M.A., 2007. Imaging mass spectrometry. Mass Spectrom. Rev. 26, 606–643. https://doi.org/10.1002/mas.20124

McFadden, G.I., 1995. Chapter 12 In Situ Hybridization. Methods Cell Biol. 49, 165–183. https://doi.org/10.1016/S0091-679X(08)61453-3

McFarlane HE, Shin JJH, Bird DA, & Samuels AL (2010) Arabidopsis ABCG Transporters, Which Are Required for Export of Diverse Cuticular Lipids, Dimerize in Different Combinations. The Plant Cell 22(9):3066-3075.

Meihls, L.N., Handrick, V., Glauser, G., Barbier, H., Kaur, H., Haribal, M.M., Lipka, A.E., Gershenzon, J., Buckler, E.S., Erb, M., Kollner, T.G., Jander, G., 2013. Natural variation in maize aphid resistance is associated with 2,4-dihydroxy-7-methoxy- 1,4-benzoxazin-3-one glucoside methyltransferase activity. Plant Cell 25, 2341. https://doi.org/10.1105/tpc.113.112409

Moore, T.S. ed (1993) Lipid Metabolism in Plants. Boca Raton: CRC Pres

Nelson, T., Langdale, J.A., 1989. Patterns of Leaf Development in C4 Plants. Plant Cell 1, 3–13. https://doi.org/10.2307/3869057

Nemes, P., Vertes, A., 2007. Laser Ablation Electrospray Ionization for Atmospheric Pressure, in Vivo, and Imaging Mass Spectrometry. Anal. Chem. 79, 8098–8106. https://doi.org/10.1021/ac071181r

Ow, D.W., De Wet, J.R., Helinski, D.R., Howell, S.H., Wood, K. V, Deluca, M., 1986. Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. Science 234, 856. https://doi.org/10.1126/science.234.4778.856

Pascal, S., Bernard, A., Sorel, M., Pervent, M., Vile, D., Haslam, R.P., Napier, J.A., Lessire, R., Domergue, F., Joubès, J., 2013. The Arabidopsis cer26 mutant, like the cer2 mutant, is specifically affected in the very long chain fatty acid elongation process. Plant J. 73, 733–746. https://doi.org/10.1111/tpj.12060

Perera, M.A.D.N., Qin, W., Yandeau-Nelson, M., Fan, L., Dixon, P., Nikolau, B.J., 2010. Biological origins of normal-chain hydrocarbons: A pathway model based on cuticular wax analyses of maize silks. Plant J. 64, 618–632. https://doi.org/10.1111/j.1365-313X.2010.04355.x

Pighin JA, et al. (2004) Plant Cuticular Lipid Export Requires an ABC Transporter. Science 306(5696):702-704.

Post-Beittenmiller, D., 1996. Biochemistry and Molecular Biology of Wax Production in Plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 405–430. https://doi.org/10.1146/annurev.arplant.47.1.405

Riederer, M., 2006. Thermodynamics of the water permeability of plant cuticles : characterization of the polar pathway 1–6. https://doi.org/10.1093/jxb/erl053

Sage, R.F. and Monson, R.K. eds (1999) C4 Plant Biology. San Diego: Academic Press

Schnable PS, Stinard PS, Wen TJ, Heinen S, Weber D, Zhang L, Hansen JD, and Nikolau BJ (1994) The genetics of cuticular wax biosynthesis, Maydica 39, 279-287

Shroff, R., Schramm, K., Jeschke, V., Nemes, P., Vertes, A., Gershenzon, J., Svatoš, A., 2015. Quantification of plant surface metabolites by matrix-assisted laser desorption-ionization mass spectrometry imaging: Glucosinolates on Arabidopsis thaliana leaves. Plant J. 81, 961–972. https://doi.org/10.1111/tpj.12760

St-Pierre, B., Luca, V. De, 2000. Chapter Nine Evolution of acyltransferase genes: Origin and diversification fo the BAHD superfamily of acyltransferases involved in secondary metabolism. Recent Adv. Phytochem. 34, 285–315. https://doi.org/10.1016/S0079-9920(00)80010-6

Svatoš, A., 2010. Mass spectrometric imaging of small molecules. Trends Biotechnol. 28, 425–434. https://doi.org/10.1016/j.tibtech.2010.05.005

Tacke, E., Korfhage, C., Michel, D., Maddaloni, M., Motto, M., Lanzini, S., Salamini, F., Döring, H.-P., 1995. Transposon tagging of the maize Glossy2 locus with the transposable element En/Spm. Plant J. 8, 907–917. https://doi.org/10.1046/j.1365-313X.1995.8060907.x

Tulloch, A.P. (1976) Biochemistry of plant waxes, in P.E. Kolattukudy (ed.), Chemistry and Biochemistry of Natural Waxes, Elsevier Press, New York, pp. 235 – 287

Vogg, G., Fischer, S., Leide, J., Emmanuel, E., Jetter, R., Levy, A.A., Riederer, M., 2004. Tomato fruit cuticular waxes and their effects on transpiration barrier properties: functional

characterization of a mutant deficient in a very-long-chain fatty acid β-ketoacyl-CoA synthase. J. Exp. Bot. 55, 1401–1410.

von Caemmerer, S. and Furbank, R.T. (2003) The C4 pathway: an efficient CO2 pump. 796 Photosynthesis Res., 77, 191-207.

Wang, X., Guan, Y., Zhang, D., Dong, X., Tian, L., Qu, L.Q., 2017. A β -Ketoacyl-CoA Synthase Is Involved in Rice Leaf Cuticular Wax Synthesis and Requires a CER2-LIKE Protein as a Cofactor. Plant Physiol. 173, 944–955. https://doi.org/10.1104/pp.16.01527

Woo, K.C., Anderson, J.M., Boardman, N.K., Downton, W.J.S., Osmond, C.B. and Thorne, 815 S.W. (1970) Deficient Photosystem II in Agranal Bundle Sheath Chloroplasts of C(4) 816 Plants. Proc. Natl. Acad. Sci. U. S. A., 67, 18-25.

Xia, Y., Nikolau, B.J., Schnable, P.S., 1996. Cloning and characterization of CER2, an Arabidopsis gene that affects cuticular wax accumulation. Plant Cell 8, 1291–1304. https://doi.org/10.1105/tpc.8.8.1291

Xu, C., Fan, J., Cornish, A.J., Benning, C., 2008. Lipid trafficking between the endoplasmic reticulum and the plastid in Arabidopsis requires the extraplastidic TGD4 protein.(RESEARCH ARTICLES). Plant Cell 20, 2190. https://doi.org/10.1105/tpc.108.061176

Yu, D., Ranathunge, K., Huang, H., Pei, Z., Franke, R., Schreiber, L., He, C., 2008. Wax Crystal-Sparse Leaf1 encodes a β–ketoacyl CoA synthase involved in biosynthesis of cuticular waxes on rice leaf. Planta 228, 675–685. https://doi.org/10.1007/s00425-008-0770-9

Zavalin, A., Yang, J., Hayden, K., Vestal, M., Caprioli, R.M., 2015. Tissue protein imaging at 1 μm laser spot diameter for high spatial resolution and high imaging speed using transmission geometry MALDI TOF MS. Anal. Bioanal. Chem. 407, 2337–2342. https://doi.org/10.1007/s00216-015-8532-6

APPENDIX A. HETEROLOGOUS CO-EXPRESSION OF MAIZE KCS GENES WITH GLOSSY2 AND GLOSSY2-LIKE MUTATED AT THE BAHD CALYTIC MOTIF IN SACCHAROMYCES CEREVISIAE

Introduction

GLOSSY2 and GLOSSY2-LIKE proteins are classified as BAHD acyltransferases based on their protein sequence where they possess the "HXXXDX" catalytic motif, however lack the second "DFGWG" BAHD motif. Both these proteins are classified in the sub-clade-II of BAHD enzymes along with CER2 proteins from Arabidopsis and rice (10,11). The BAHD acyl transferase family derived its name from the first four biochemically characterized enzymes, BEAT (benzyl alcohol acetyl-transferase, (1)), AHCT (anthocyanin-O-hydroxycinnamoyl transferase, (2)), HCBT (anthranilate-N-hydroxycinnamoyl/ benzoyl transferase, (3)) and DAT (deacetylvindoline 4-O-acetyl-transferase, (4)). Several BAHD acyltransferase enzymes have been enzymatically characterized and crystal structures have been solved for a number of plant BAHD acyltransferases which include vinorine synthase from *Rauvolfia* (5), anthocyanin malonyl transferase from *Chrystanthemum* (6), malonyltransferse from *Nicotiana* tabacum (7) and hydroxylcinnamoyl transferases from *Coffea canefora* and *Sorghum bicolor* (8,9). Recent yeast expression studies on rice CER2 demonstrated that this protein interacts with a KCS (WSL4) involved in the first step of fatty acid elongation, and this interaction is dependent on the BAHD catalytic motif of OsCER2 (10). *In planta* studies demonstrated that GL2 and GL2-LIKE behave differently in their BAHD enzyme-type reactions, where the latter was indicative of a BAHD-type function (Alexander, 2019; this thesis). Therefore, using yeast heterologous systems we investigated if mutation of the BAHD catalytic motif had an effect on the GL2 and GL2-LIKE when co-expressed with maize KCS.

Result and Discussion

Using site-directed mutagenesis, the His- and Asp- residues of the HXXXD BAHD catalytic domain of GL2 and GL2-LIKE were mutated to Ala, inorder to determine if the *Gl2* homologs altered the very long chain fatty acid profiles in yeast when co-expressed with two maize KCS; *ZmKCS5* (AC233893.1_FGT003) and *ZmKCS6* (GRMZM2G164974). Co-expression was done in *elo2 elo3* double mutant haploid strain that are synthetically lethal but can be rescued by the episomal expression of ZmKCS homologs (Campbell et al., 2019).

Strains expressing *ZmKCS5, ZmKCS6,* and those co-expressing *ZmKCS5* or *ZmKCS6* with either of the two *ZmGl2* homologs or the two *ZmGl2* homologs carrying mutations at the His- or Asp- residues of the HXXXD BAHD motif do not produce additional changes in the VLCFA profiles generated (Figure A1 and Figure A2). However, co-expression of *Gl2* homologs with ZmKCS6 resulted in changes within the existing VLCFA profiles. Specifically, the co-expression of *Gl2* with *ZmKCS6* showed decreased amounts of 26:0, 28:0 and 30:0 in comparison to *ZmKCS6* controls. The co-expression of *Gl2* carrying mutations at the His169- or Asp173- residue with *ZmKCS6* did not alter the VLCFA profiles and continues to show decreased levels of 26:0 to 30:0 (Figure A2). This indicates that *Gl2* is not dependent on the BAHD catalytic domain, however the decrease in existing VLCFA may be indicative of a negative regulation.

The co-expression of *ZmKCS6* with *Gl2-like* did not alter the existing VLCFA profiles, however co-expression of *ZmKCS6* with *Gl2-like* mutated at the His185- residue resulted in ~ 32% decrease of 26:0 in comparison to *Gl2-like* and *ZmKCS6* controls (Figure A2), however no changes were observed for the Asp189- residue mutation nor were there any changes to longer acyl chains. It is unclear why mutations at the His185-residue of *Gl2-like* when co-expressed with *ZmKCS6,* resulted in a decrease specific to 26:0. *In planta* studies on *Glossy2* and *Glossy2-like* demonstrated

differences in their BAHD roles, where the latter required the BAHD domain for epicuticular biosynthesis. Although, the effect of His185-mutation in *Gl2-like* in yeast is specific to a reduction in 26:0, the change was minor and thereby, inconclusive of the probable BAHD enzyme. Further experiments are required to understand the function of both the GL2 and GL2-LIKE protein, such as determining biochemical assays in *E.coli*.

Materials and Method

Maize genes were cloned into high-copy episomal plasmids pAG426 (URA3), or pAG423 (HIS3) (Invitrogen, Carlsbad, CA) using either Gateway® cloning system (Invitrogen, Carlsbad, CA) or In-Fusion® cloning (Takara Bio USA, Inc., Mountain View, CA). *Glossy2* and *Glossy2-like* were expressed on pAG426 (URA) and the maize KCS were expressed on pAG423 (HIS). The genes were under the control of the constitutive glyceraldehyde-3-P-dehydrogenase (GPD) promoter. Mutation of the His- and Asp- residue of the HXXXD motif of GLOSSY2 and GLOSSY2-LIKE were generated using QuikChange XL Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA). All shuttle vectors were confirmed by DNA sequencing. Yeast transformation, growth conditions, fatty acid analysis were performed as previously described in Chapter 7.

References

1. Dudareva N, D'Auria JC, Nam KH, Raguso RA, Pichersky E. Acetyl-CoA:benzylalcohol acetyltransferase--an enzyme involved in floral scent production in Clarkia breweri. Plant J. 1998;14(3):297.

2. Fujiwara H, Tanaka Y, Yonekura-Sakakibara K, Fukuchi-Mizutani M, Nakao M, Fukui Y, et al. cDNA cloning, gene expression and subcellular localization of anthocyanin 5-aromatic acyltransferase from Gentiana triflora. Plant J. 1998;16(4):421.

3. Yang Q, Reinhard K, Schiltz E, Matern U. Characterization and heterologous expression of hydroxycinnamoyl/benzoyl-CoA:anthranilate N-hydroxycinnamoyl/benzoyltransferase from elicited cell cultures of carnation, Dianthus caryophyllus L. Charact heterologous Expr hydroxycinnamoyl/benzoyl-CoAanthranilate N-

hydroxycinnamoyl/benzoyltransferase from elicited cell Cult carnation, Dianthus caryophyllus L. 1997;(6):777–89.

4. St-Pierre B, Laflamme P, Alarco AM, De Luca V. The terminal O-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A-dependent acyl transfer. Plant J. 1998;14(6):703.

5. Ma X, Koepke J, Panjikar S, Fritzsch GG, Stöckigt J. Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. J Biol Chem. 2005;280(14):13576–83.

6. Unno H, Ichimaida F, Suzuki H, Takahashi S, Tanaka Y, Saito A, et al. Structural and mutational studies of anthocyanin malonyltransferases establish the features of BAHD enzyme catalysis. J Biol Chem. 2007;282(21):15812–22.

7. Manjasetty B, Yu X-H, Panjikar S, Taguchi G, Chance M, Liu C-J. Structural basis for modification of flavonol and naphthol glucoconjugates by Nicotiana tabacum malonyltransferase (Nt MaT1). An Int J Plant Biol. Berlin/Heidelberg; 2012;236(3):781–93.

8. Lallemand LA, McCarthy JG, McSweeney S, McCarthy AA. Purification, crystallization and preliminary X-ray diffraction analysis of a hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase (HCT) from Coffea canephora involved in chlorogenic acid biosynthesis. Acta Crystallogr Sect F Struct Biol Cryst Commun. International Union of Crystallography; 2012;68(7):824–8.

9. Walker AM, Hayes RP, Youn B, Vermerris W, Sattler SE, Kang C. Elucidation of the structure and reaction mechanism of sorghum hydroxycinnamoyltransferase and its structural relationship to other coenzyme a-dependent transferases and synthases. Plant Physiol [Internet]. 2013;162(2):640–51. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3668059&tool=pmcentrez&rendertype=abstract

10. Wang X, Guan Y, Zhang D, Dong X, Tian L, Qu LQ. A β-Ketoacyl-CoA Synthase Is Involved in Rice Leaf Cuticular Wax Synthesis and Requires a CER2-LIKE Protein as a Cofactor. Plant Physiol [Internet]. 2017 Feb 1;173(2):944 LP-955. Available from: http://www.plantphysiol.org/content/173/2/944.abstract

11. Haslam TM, Haslam R, Thoraval D, Pascal S, Delude C, Domergue F, et al. ECERIFERUM2-LIKE Proteins Have Unique Biochemical and Physiological Functions in Very-Long-Chain Fatty Acid Elongation. Plant Physiol [Internet]. 2015;167(3):682–92. Available from: http://www.plantphysiol.org/content/early/2015/01/16/pp.114.253195.abstract?papetoc

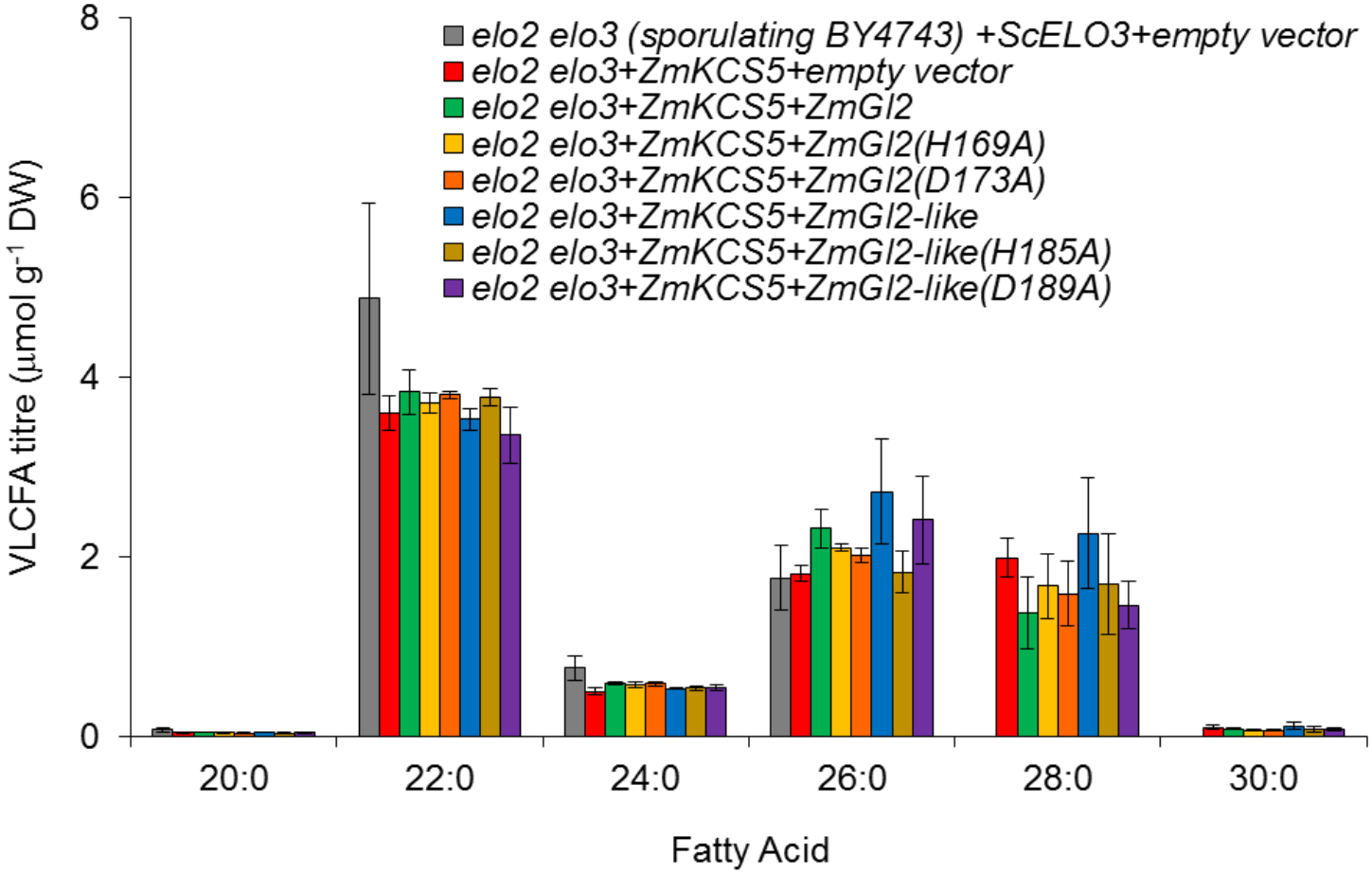


Figure A1. Total of VLCFA accumulation of yeast *elo2elo3* strains co-expressing maize KCS5 with *Glossy2* gene family or *Glossy2* gene family mutated at their catalytic HXXXD motif, under GPD inducible promoter, n=3 replicates per genotype.

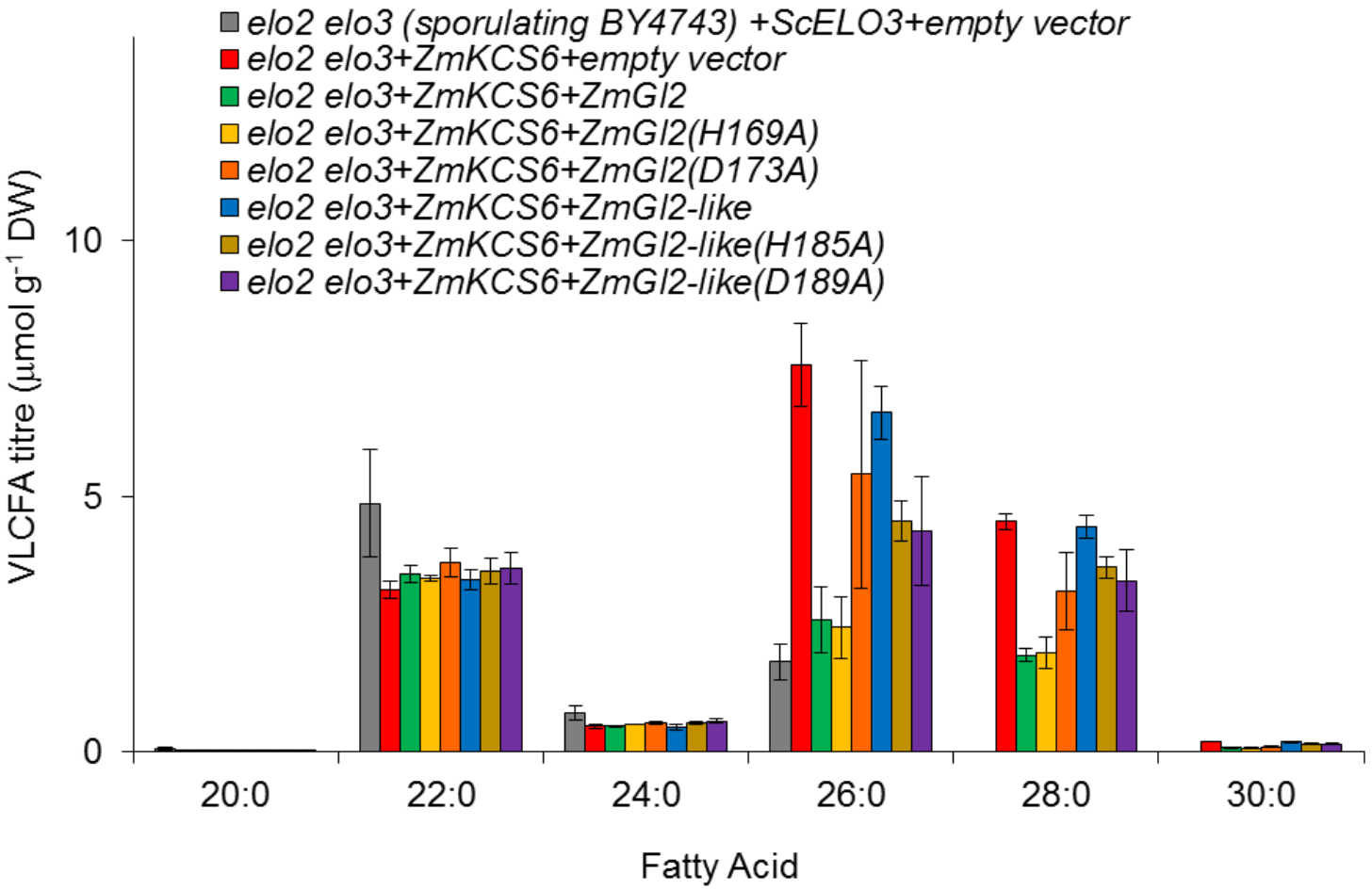


Figure A2. Total VLCFA accumulation of yeast *elo2elo3* strains co-expressing maize KCS6 with *Glossy2* gene family or *Glossy2* gene family mutated at their catalytic HXXXD motif, under GPD inducible promoter, n=3 replicates per genotype.

APPENDIX B. HETEROLOGOUS EXPRESSION OF GLOSSY2 IN E. COLI

Introduction

The gap that exist in our understanding of the mechanism of action of the GLOSSY2 protein has made it even more important to understand the structure of this protein and develop an *in vitro* assay that is representative of its *in vivo* function. For example, the GLOSSY2 (and its close homolog CER2 from Arabidopsis) belong to the BAHD superfamily of enzymes that catalyze acyl-CoA dependent acyltransferase reaction that generate amides and esters. However, recent *in vivo* genetic based correlative studies have been interpreted as indicating a role for these proteins in the fatty acid elongation process (1,2). However, there is no apparent role for an acyltransferase-type reaction based on current knowledge of the FAE mechanism that minimally requires iterations of 4 reactions: a) Claisen condensation; b) 3-ketoacyl reduction; c) 3-hydroxyacyl dehydration and d) enoyl reduction. Indeed, to date the specific function of none of the Clade II BAHD proteins, which is defined by CER2 and GLOSSY2 proteins, have been determined. In light of this dilemma, a system was developed to express and purify the GLOSSY2 protein in sufficient quantities to enable *in vitro* characterizations that can be used to explore structure and function of the protein, which can provide insights concerning its *in vivo* function.

Results and Discussion

The GLOSSY2 protein was recombinantly expressed with an N-terminal His-tag using a pDEST17 expression vector, in *E. coli* strain C43. The presence of the His-tag facilitated purification of the recombinant protein via Ni-NTA affinity chromatography. This resulted in pure preparations of GLOSSY2, with a typical yield of 6 mg of protein per litre of culture. SDS-PAGE analysis of establish the purity of these preparations and the GLOSSY2 protein migrates as a

singular protein band with an apparent molecular weight of 46 kDa, in very good agreement with the sequence based prediction (Figure B1).

Gel filtration chromatography on a Superdex 200 10/300 GL (GE Healthcare Life Sciences, Pittsburg, PA) column was used to characterize the native form of the purified GLOSSY2 protein. These analyses indicate that this protein appears to be multimeric that is assembled by disulfide linkages. Thus, in the presence of the reducing agent, TCEP, the protein elutes with an apparent molecular weight of 39 $\pm$ 0.9 kDa, which is the monomer. However, in the absence of TCEP, the protein elutes as 4 distinct peaks, with apparent molecular weights of 200 $\pm$ 9.7 kDa, 135 $\pm$ 6.5 kDa, 79 $\pm$ 4.6 kDa, and 39 $\pm$ 1.3 kDa, which we infer corresponds to tetrameric, trimeric, dimeric and monomeric forms of the protein, respectively (Figure B2).

Circular Dichrosim spectra indicate that both the oligomeric and monomeric GL2 protein preparations are folded (Figure B3). To begin the process of identifying potential substrates or ligands that may interact with the GL2 protein, we applied the general principle that upon the binding of such molecules, protein structure often coalesces around the bound substrate/ligand, which affects the thermal stability of the protein. Thus, we employed fluorescence-based thermal shift assays (3,4) on both monomeric and oligomeric preparations of the GL2 protein, to survey the ability of the protein to interact with different acyl CoA molecules, which are the acyl substrates of BAHD enzymes (5,6).

These thermal shift assays were primarily developed for high-throughput drug discovery (3,7,8) and have more recently been used to optimize conditions for protein stabilization and crystallization (4,9). The assay measures the changes in protein melting temperature (T_m) upon addition of additives (such as ligands, buffers, salts) by use of fluorescent hydrophobic dyes which bind to the exposed hydrophobic areas of the melted protein. Shifts in the melting temperature can

either be positive, indicative of binding and stabilization by the ligand/substrate or negative suggesting a destabilization of the protein by the ligand/substrate.

The ligands tested with GL2 included free CoA and a collection of acyl CoAs of varying acyl-chain lengths (2:0, 3:0, 16:0, 18:0, and 26:0). Ligand controls, no-protein controls, and proteins without ligands were run as controls for these experiment. The melting temperature for both the monomeric and oligomeric GL2 protein is 48.6 $\pm$ 0.4^{o}C. All the ligands tested decreased the melting temperature of the protein by 2.6 $\pm$ 0.4^{o}C indicating a destabilizing effect (Figure B4). These results demonstrate that GL2 is capable of interacting with the tested acyl-CoAs, but these interactions destabilize the protein to increasing temperature. The destabilizing effect is difficult to explain without additional structural information, and may in fact be due to these acyl-CoA ligands acting as detergents, a particular characteristic of the longer chain acyl-CoA (e.g., 26:0-CoA). One interesting observation that may be of significance, is the concentration dependent destabilization of GL2 by 16:0-CoA. This is in contrast to all other acyl-CoAs tested, and in particular is different behavior than obtained with 18:0-CoA, which differs by just 2-carbon atoms. Moreover this effect is specific for the monomeric form of the GL2 protein, and is not observed with the oligomeric form, which maybe an indication that the interaction between GL2 and 16:0-acyl-CoA is of functional significance.

Materials and Methods

Gene cloning

The ORFs of *Glossy2* (GRMZM2G098239) was codon-optimized for yeast using GeneOptimizer (GeneArt, LifeTechnologies) and cloned into pDEST17 vector using Gateway cloning (Invitrogen, Carlsbad, CA), resulting in the pDEST_GLOSSY2 construct that carried an

N-terminal His-tagged *Glossy2* gene sequence. The resulting plasmid was confirmed by sequencing.

Expression and purification of recombinant GLOSSY2 protein

E.coli OverExpress ™ C43 (Lucigen, Middletown, WI) strain was used for the expression of GLOSSY2 protein. The transformants were grown according to the manual for C43 cells. Briefly, the cells were grown overnight at 37ºC in 5 mL Luria-Bertani medium supplemented with 100 μg/ml ampicillin (Research Products International Corps., Mount Prospect, IL) and 0.2% (w/v) glucose (to minimize amount of expression of target protein prior to induction). The pre-culture was then transferred into 500 mL LB medium supplemented with 100 μg/ml ampicillin and induced by the addition of 0.5mM IPTG (Gold Biotechnology, Olivette, MO) when the OD_{600} of the culture was between 0.8-1.0.

After incubation for another 16-18 h at 22ºC, cells were harvested by centrifugation (5000 rpm, 4ºC, 15 min). Soluble protein was extracted by first suspending the cell pellet in lysis buffer (4 mL per g wet weight; 50mM Potassium phosphate buffer pH 7.8, 150 mM NaCl, 50 mM KCl, 10 mM imidazole, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride, 0.5% Triton-X 100, cOmplete™, EDTA-free protease inhibitor cocktail (Roche Applied Science, Indianapolis, IN)), followed by sonication on ice (10 x 15s pulses separated by 1 min intervals), and centrifugation (18,000 x g, 4ºC, 30 min). The soluble protein fraction was filtered through a 0.45 μm syringe filter (Corning Inc., Corning, NY) and applied thrice to a 5 mL Ni-NTA His-bind resin (PerfectPro™ Ni-NTA Agarose; Fisher Scientific, Pittsburg, PA) that has been equilibrated with histidine binding buffer (0.3 M Nacl, 50 mM Sodium phosphare, 10 mM imidazole; 5 column volume). After washing the unbound protein with 20 column volumes of wash buffers I and II (0.3 M NaCl, 50 mM Sodium phosphate) supplemented with 20 mM and 50 mM imidazole

respectively, the bound protein was eluted with 10 column volume of the same buffer containing 200 mM imidazole. The purified Hi-tagged GL2 protein was dialyzed against potassium phosphate buffer, pH 7.8 containing 150 mM NaCl and 50 mM KCl, and concentrated by centrifugation at 3900 rpm at 4^{o}C using 10,000 Da cutoff ultracentrifugation filter (Millipore-Amicon, Billerica, MA) to a final volume of 1 ml. The concentrated protein was supplemented with 15% glycerol and stored at -80^{o}C. Protein concentrations were determined by Bradford assay (Biorad®, Hercules, CA) with bovine serum albumin (BSA) as a standard. Protein purity was assessed by Coomassie-staining SDS-PAGE gels.

Fast paced liquid chromatography (FPLC) analysis

The molecular weight of purified protein complexes was determined by FPLC using a Superdex 200 10/300 GL gel filtration column (GE Healthcare Life Sciences, Pittsburg, PA) using a Atka FPLC system (GE Healthcare Bio-Sciences, Pittsburg, PA). Typically a 100 μL loop was used to inject 3 mg of protein. 50mM potassium phosphate buffer pH 7.8, containing 150 mM NaCl, 50 mM KCl, and 10% glycerol was used as the mobile phase, and maintained at a flowrate of 0.4 mL/min. The column was equilibrated in the mobile phase for 2-h prior to protein injection. Protein was monitored in-line using a UV detector at 280nm and FPLC fractions were collected in 400 μL increments. To detect the native state of the protein, fractions were run on SDS-PAGE, and visualized via Coomassie staining. The effect of disulfide interactions in the protein was tested by adding 1 mM TECP to the mobile phase, incubated for 2-h in ice and then analyzed as described above. Molecular weights were determined using BioRad gel filtration standard (Hercules, CA)

Circular dichroism spectroscopy of GL2 protein

CD spectra of purified GL2 proteins (0.15–0.2 mg/mL in 50mM potassium phosphate buffer, pH 7.8 containing 150 mM NaCl and 50 mM KCl) were collected with Jasco J-715

Spectropolarimeter, in a 0.1cm cell at 25°C. Far UV spectra were scanned from 260 nm to 190 nm, recorded with a bandwidth of 1.0 nm and a time response of 4 s with total of 2 data accumulations. A blank was taken by diluting 50mM potassium phosphate buffer pH 7.8 containing 150 mM NaCl, 50 mM KCl, and 10% glycerol into ddH2O at the same ratio as the sample, and this blank was then subtracted from the sample to provide the ellipticity (mdeg).

Thermal shift assays

Differential scanning fluorimetry (DSF) was determined by protein thermal shift assays in a 96-well plate setup using a StepOnePlus™ Real-Time PCR System (Applied Biosystems life technologies, Carlsbad, CA). Data was logged using StepOne™ Software (Applied Biosystems life technologies, Carlsbad, CA). Protein unfolding was monitored using the dye SYPRO® Orange (Applied Biosystems life technologies, Carlsbad, CA). Optimization of protein to dye ratios were performed in potassium phosphate buffer pH 7.8 (50 mM potassium phosphate buffer pH 7.8 containing 150 mM NaCl, 50 mM KCl, and 10% glycerol) with stock SYPRO® Orange 5000X concentration diluted in ddH2O to a working concentration of 50X. Protein concentration of 1mg/mL and final dye concentration of 5X were used per 20 μL reaction. The temperature was varied from 25^{o}C to 99^{o}C at a rate of 2^{o}C per minute. SYPRO® Orange fluorescence was measured at an excitation and emission wavelengths of 492 nm and 610 nm respectively. The melting temperature (Tm) values were estimated using the Boltzmann method in the Protein Thermal Shift™ Software v1.0 (Applied Biosystems life technologies, Carlsbad, CA). Various concentrations of CoA ligands were tested using the optimum protein to dye ratio. The change in Tm (ΔTm) was estimated by the difference in melting temperature between ligand-bound and ligand-free protein.

References

1. Bianchi. Glossy mutants of maize. VI. Chemical constituents of glossy-2 epicuticular waxes. Maydica [Internet]. 1975 Jan 1;20(4):165–73. Available from: https://eurekamag.com/research/000/389/000389438.php

2. Haslam TM, Haslam R, Thoraval D, Pascal S, Delude C, Domergue F, et al. ECERIFERUM2-LIKE Proteins Have Unique Biochemical and Physiological Functions in Very-Long-Chain Fatty Acid Elongation. Plant Physiol [Internet]. 2015;167(3):682–92. Available from: http://www.plantphysiol.org/content/early/2015/01/16/pp.114.253195.abstract?papetoc

3. Ericsson UB, Hallberg BM, DeTitta GT, Dekker N, Nordlund P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. Anal Biochem. 2006;357(2):289–98.

4. Niesen FH, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat Protoc [Internet]. 2007;2(9):2212–21. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=17853878&retmode=ref&cmd=prlinks

5. St-Pierre B, Luca V De. Chapter Nine Evolution of acyltransferase genes: Origin and diversification fo the BAHD superfamily of acyltransferases involved in secondary metabolism. Recent Adv Phytochem. 2000;34(C):285–315.

6. D'Auria JC. Acyltransferases in plants: a good time to be BAHD. Curr Opin Plant Biol. 2006;9(3):331–40.

7. Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, Graf E, et al. High-Density Miniaturized Thermal Shift Assays as a General Strategy for Drug Discovery. J Biomol Screen. 2001;6(6):429–40.

8. Lo M-C, Aulabaugh A, Jin G, Cowling R, Bard J, Malamas M, et al. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. Anal Biochem [Internet]. Academic Press; 2004 Sep 1 [cited 2019 Jan 13];332(1):153–9. Available from: https://www.sciencedirect.com/science/article/pii/S0003269704003756?via%3Dihub

9. Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ, Wasney GA, et al. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proc Natl Acad Sci [Internet]. National Academy of Sciences; 2006 Oct 24 [cited 2019 Jan 14];103(43):15835–40. Available from: https://www.pnas.org/content/103/43/15835

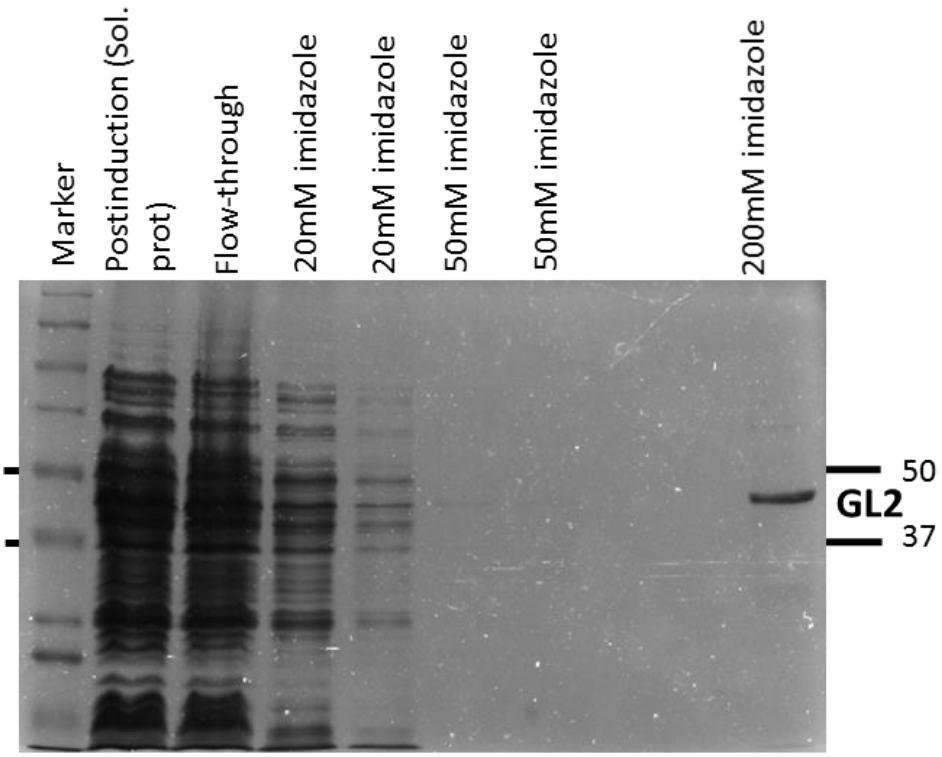
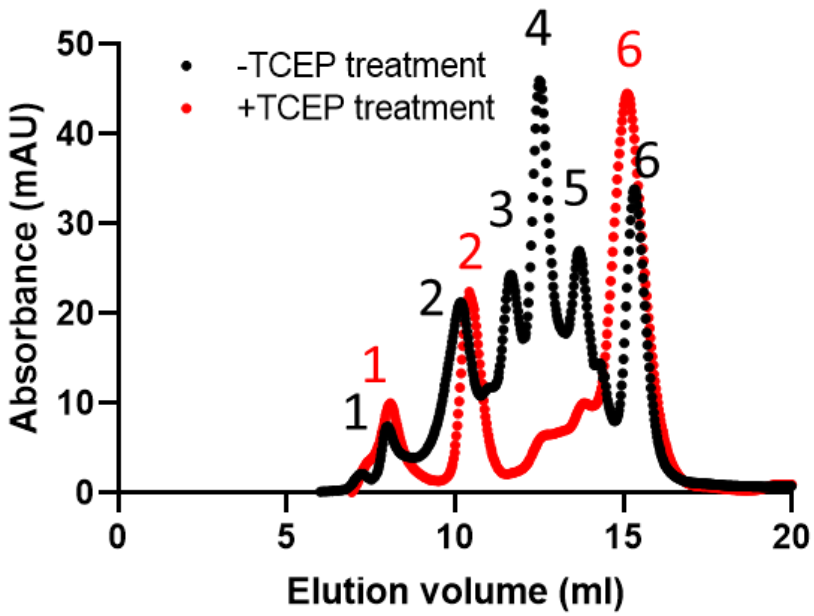


Figure B1. SDS-PAGE analysis of purified, recombinantly expressed GLOSSY2 protein visualized by staining with Coomassie Brilliant Blue. GL2 is the 46-kDa protein band recovered in the sample eluted from the Ni-affinity chromatography column with buffer containing 200 mM imidazole.



Peak #	Calculated Molecular Weight (kDa)
3	200 ± 9.7
4	135 ± 6.5
5	79 ± 4.6
6	39 ± 1.3

Figure B2. Size-exclusion FPLC chromatograms of GLOSSY2 protein before (black) and after (red) treatment with 1 mM of the reducing agent, TCEP and their calculated mass denoted in the table, is averaged between two independent GLOSSY2 purifications.

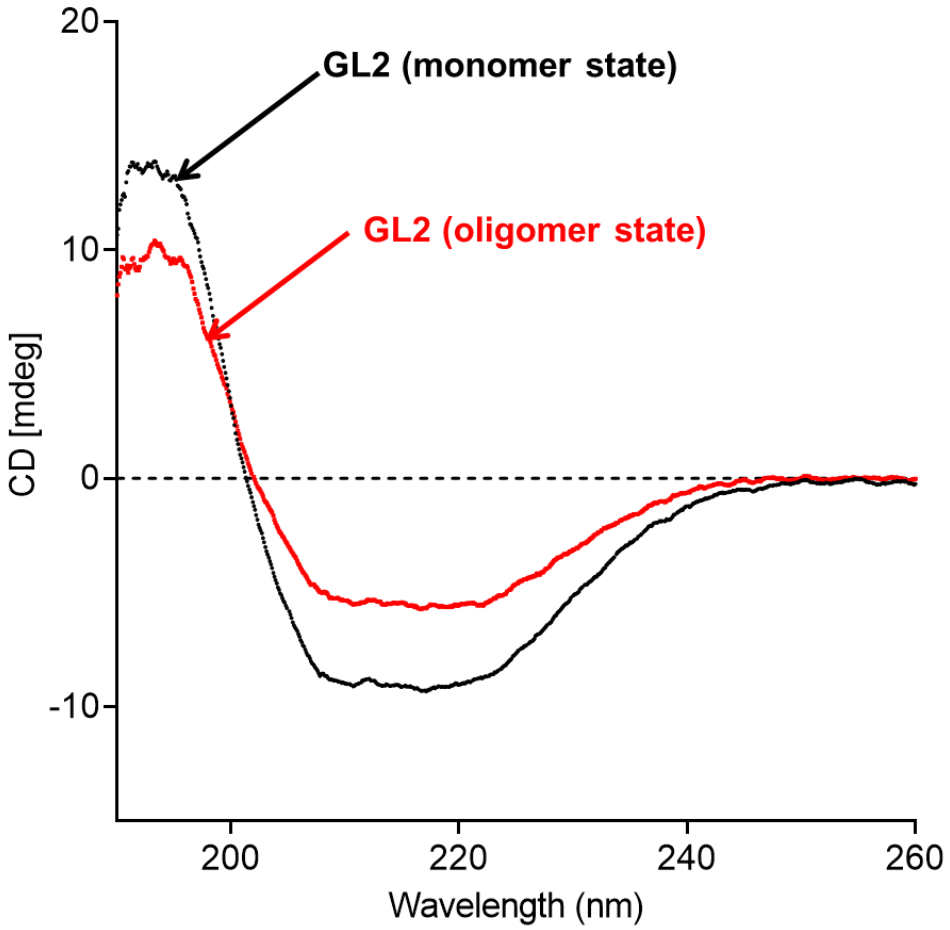


Figure B3. Comparison of CD spectra of monomeric (black) and oligomeric (red) forms of the purified GLOSSY2 protein. Measurements are an average of three scans with a reading every 0.1 nm.

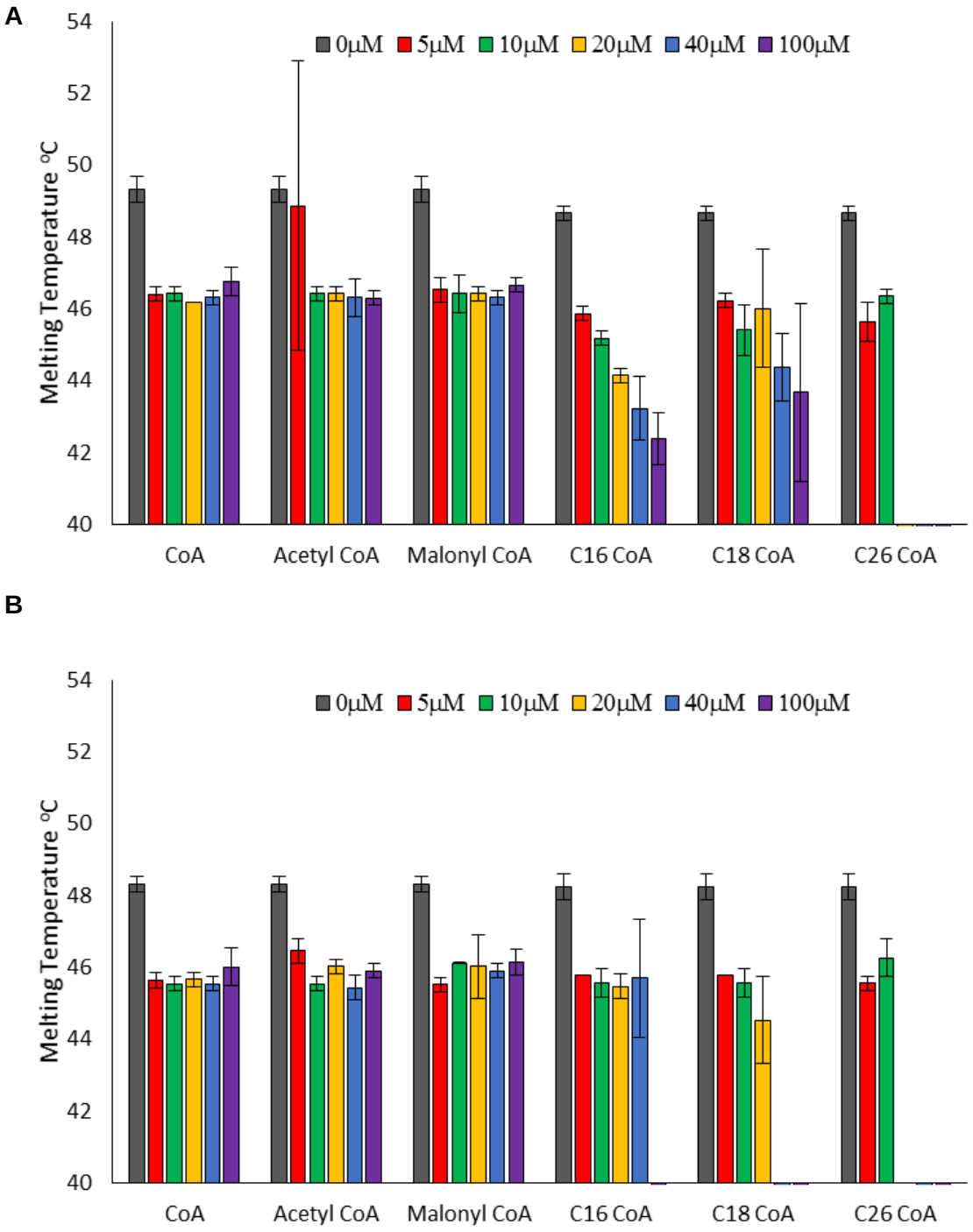


Figure B4. Thermal stabilization of the monomeric (A) and oligomeric (B) forms of GLOSSY2 protein by different acyl-CoAs, determined via fluorescence based thermal shift assay.